

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2002, 02:19:50 ; Search time 2768 Seconds
(without alignments)
3690.421 Million cell updates/sec

Title: US09897438BK-2
Perfect score: 351
Sequence: 1 GARCARTGYGNACNATNAT.....AYCARGCNTGTGGCNYTN 351

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 50.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rod:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	16	4.6	23	6	AR136101	Sequence AR136101 Sequence
2	16	4.6	885	6	AX047947	Sequence AX047947 Sequence
3	16	4.6	1044	6	AX047954	Sequence AX047954 Sequence
4	16	4.6	2379	6	AX211531	Sequence AX211531 Sequence
5	15	4.3	426	6	AX055062	Sequence AX055062 Sequence
6	15	4.3	696	6	AX255037	Sequence AX255037 Sequence
7	15	4.3	729	6	AX255040	Sequence AX255040 Sequence
8	15	4.3	987	6	AX068381	Sequence AX068381 Sequence
9	15	4.3	1155	6	AX108644	Sequence AX108644 Sequence
10	15	4.3	2019	6	AX179624	Sequence AX179624 Sequence
11	15	4.3	2277	6	AR004980	Sequence AR004980 Sequence
12	14	4.0	20	6	I26713	Sequence I26713 Sequence 14
13	14	4.0	1026	6	AX350993	Sequence AX350993 Sequence
14	14	4.0	1134	6	AX172419	Sequence AX172419 Sequence
15	14	4.0	1584	6	AX167244	Sequence AX167244 Sequence
16	14	4.0	1629	6	AX167247	Sequence AX167247 Sequence
17	14	4.0	1695	6	AX350990	Sequence AX350990 Sequence
18	14	4.0	1728	6	AX237003	Sequence AX237003 Sequence
19	14	4.0	2094	6	AX350978	Sequence AX350978 Sequence
20	14	4.0	2094	6	AX360312	Sequence AX360312 Sequence
21	13	3.7	20	6	A23188	Sequence A23188 Artificial
22	13	3.7	20	6	AR050672	Sequence AR050672 Sequence
23	13	3.7	20	6	E08318	Sequence E08318 Probe for d
24	13	3.7	23	6	AR009221	Sequence AR009221 Sequence
25	13	3.7	23	6	I49042	Sequence I49042 Sequence 13
26	13	3.7	23	6	I73415	Sequence I73415 Sequence 13
27	13	3.7	51	6	AX356144	Sequence AX356144 Sequence
28	13	3.7	72	6	AR042415	Sequence AR042415 Sequence
29	13	3.7	351	6	AX154583	Sequence AX154583 Sequence
30	13	3.7	390	6	AX193710	Sequence AX193710 Sequence
31	13	3.7	693	6	AX151702	Sequence AX151702 Sequence
32	13	3.7	693	6	AX193707	Sequence AX193707 Sequence
33	13	3.7	696	6	AX280162	Sequence AX280162 Sequence
34	13	3.7	729	6	AX350987	Sequence AX350987 Sequence
35	13	3.7	747	6	AX193684	Sequence AX193684 Sequence
36	13	3.7	804	6	AX398327	Sequence AX398327 Sequence
37	13	3.7	882	6	AX398330	Sequence AX398330 Sequence
38	13	3.7	906	6	AX398703	Sequence AX398703 Sequence
39	13	3.7	1098	6	AX163791	Sequence AX163791 Sequence
40	13	3.7	1602	6	AR150957	Sequence AR150957 Sequence
41	13	3.7	1644	6	AX269154	Sequence AX269154 Sequence
42	13	3.7	2082	6	AR166376	Sequence AR166376 Sequence
43	13	3.7	2109	6	AX350975	Sequence AX350975 Sequence
44	13	3.7	2214	6	AX350981	Sequence AX350981 Sequence
45	13	3.7	2217	6	AX364578	Sequence AX364578 Sequence

ALIGNMENTS

RESULT 1	AR136101	Sequence 35 from patent US 6136581.	23 bp	DNA	linear	PAT 16-JUN-2001
LOCUS	AR136101	Sequence 35 from patent US 6136581.				
DEFINITION	AR136101	Sequence 35 from patent US 6136581.				
ACCESSION	AR136101	Sequence 35 from patent US 6136581.				
VERSION	AR136101.1	GI:14476773				
KEYWORDS						
SOURCE		Unknown.				
ORGANISM		Unknown.				
REFERENCE		Unclassified.				
AUTHORS		1 (bases 1 to 23)				
TITLE		Joho,K.E. and Plowman,G.D.				
JOURNAL		Kinase genes and uses				
FEATURES		Patent: US 6136581-A 35 24-OCT-2000;				
		Location/Qualifiers				

```
source 1. .23
/organism="unknown"
BASE COUNT 3 a 1 c 8 g 4 t 7 others
ORIGIN

Query Match 4.6%; Score 16; DB 6; Length 23;
Best Local Similarity 100.0%; Pred. No. 7.3e-21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 322 GCGARGTNTAYGARG 337
|||||
Db 7 GCGARGTNTAYGARG 22

RESULT 2
AX047947
LOCUS AX047947 885 bp DNA linear PAT 15-DEC-2000
DEFINITION Sequence 14 from Patent WO0070045.
ACCESSION AX047947
VERSION AX047947.1 GI:11876870
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified
REFERENCE 1 (bases 1 to 885)
AUTHORS Barclay,A.N., Brown,M.H., Gorman,D.M., Lanier,L.L., Wright,G.J.,
Cherwinski,H., Phillips,J.H., Hoek,R.M. and Sedgwick,J.D.
TITLE Ox2 receptor homologs
JOURNAL Patent: WO 0070045-A 14 23-NOV-2000;
MEDICAL RESEARCH COUNCIL (GB) ; SCHERING CORPORATION (US)
FEATURES
source Location/Qualifiers
1. .885
/db_xref="taxon:32644"
BASE COUNT 182 a 98 c 133 g 114 t 358 others
ORIGIN

Query Match 4.6%; Score 16; DB 6; Length 885;
Best Local Similarity 100.0%; Pred. No. 1.5e-21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 229 AAYGTNWSNACNGTNA 244
|||||
Db 697 AAYGTNWSNACNGTNA 712

RESULT 3
AX047954
LOCUS AX047954 1044 bp DNA linear PAT 15-DEC-2000
DEFINITION Sequence 21 from Patent WO0070045.
ACCESSION AX047954
VERSION AX047954.1 GI:11876877
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1044)
AUTHORS Barclay,A.N., Brown,M.H., Gorman,D.M., Lanier,L.L., Wright,G.J.,
Cherwinski,H., Phillips,J.H., Hoek,R.M. and Sedgwick,J.D.
TITLE Ox2 receptor homologs
JOURNAL Patent: WO 0070045-A 21 23-NOV-2000;
MEDICAL RESEARCH COUNCIL (GB) ; SCHERING CORPORATION (US)
FEATURES
source Location/Qualifiers
1. .1044
/db_xref="taxon:32644"
BASE COUNT 220 a 117 c 154 g 131 t 422 others
ORIGIN

Query Match 4.6%; Score 16; DB 6; Length 1044;
Best Local Similarity 100.0%; Pred. No. 1.4e-21;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 229 AAYGTNWSNACNGTNA 244
|||||
Db 697 AAYGTNWSNACNGTNA 712

RESULT 4
AX211531
LOCUS AX211531 2379 bp DNA linear PAT 07-SEP-2001
DEFINITION Sequence 12 from Patent WO0159112.
ACCESSION AX211531
VERSION AX211531.1 GI:15523810
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 2379)
AUTHORS Holloway,J.L. and Sheppard,P.O.
TITLE Anti-angiogenic intestinal peptides, z dint5
JOURNAL Patent: WO 0159112-A 12 16-AUG-2001;
ZymoGenetics, Inc. (US)
FEATURES
source Location/Qualifiers
1. .2379
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 270 a 286 c 502 g 272 t 1049 others
ORIGIN

Query Match 4.6%; Score 16; DB 6; Length 2379;
Best Local Similarity 100.0%; Pred. No. 9.7e-22;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 NWSNGGNWSNTGYMGN 138
|||||
Db 1512 NWSNGGNWSNTGYMGN 1527

RESULT 5
AX055062
LOCUS AX055062 426 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 3 from Patent WO0073458.
ACCESSION AX055062
VERSION AX055062.1 GI:12228387
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 426)
AUTHORS Conklin,D.C.
TITLE Secreted alpha-helical protein-31
JOURNAL Patent: WO 0073458-A 3 07-DEC-2000;
ZymoGenetics, Inc. (US)
FEATURES
source Location/Qualifiers
1. .426
/organism="synthetic construct"
/db_xref="taxon:32630"
misc_feature 1..426
/notes="Degenerate polynucleotide sequence of zalpha31"
/notes="n = A,T,C or G"
BASE COUNT 101 a 41 c 70 g 43 t 171 others
ORIGIN

Query Match 4.3%; Score 15; DB 6; Length 426;
Best Local Similarity 100.0%; Pred. No. 7e-19;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 263 CNGARGARGCNAARG 277
|||||
Db 77 CNGARGARGCNAARG 91

RESULT 6
AX255037
```

LOCUS AX255037 696 bp DNA linear PAT 10-OCT-2001
DEFINITION Sequence 3 from Patent WO0170986.
ACCESSION AX255037
VERSION AX255037.1 GI:16074534
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 696)
AUTHORS Konlin,D.C. and Presnell,S.R.
TITLE Helical protein zalpha51
JOURNAL Patent: WO 0170986-A 3 27-SEP-2001;
ZymoGenetics, Inc. (US)
FEATURES
source Location/Qualifiers
1..696
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="degenerate sequence"
BASE COUNT 74 a 82 c 121 g 93 t 326 others
ORIGIN
Query Match 4.3%; Score 15; DB 6; Length 696;
Best Local Similarity 100.0%; Pred. No. 5.7e-19;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 257 AYTNCNGARGARG 271
|||||
Db 449 AYTNCNGARGARG 463
RESULT 7
AX255040
LOCUS AX255040 729 bp DNA linear PAT 10-OCT-2001
DEFINITION Sequence 6 from Patent WO0170986.
ACCESSION AX255040
VERSION AX255040.1 GI:16074537
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 729)
AUTHORS Konlin,D.C. and Presnell,S.R.
TITLE Helical protein zalpha51
JOURNAL Patent: WO 0170986-A 6 27-SEP-2001;
ZymoGenetics, Inc. (US)
FEATURES
source Location/Qualifiers
1..729
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="degenerate sequence"
BASE COUNT 78 a 85 c 133 g 96 t 337 others
ORIGIN
Query Match 4.3%; Score 15; DB 6; Length 729;
Best Local Similarity 100.0%; Pred. No. 5.6e-19;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 257 AYTNCNGARGARG 271
|||||
Db 482 AYTNCNGARGARG 496
RESULT 8
AX068381
LOCUS AX068381 987 bp DNA linear PAT 25-JAN-2001
DEFINITION Sequence 4 from Patent WO0102565.
ACCESSION AX068381
VERSION AX068381.1 GI:12578542
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 987)

AUTHORS Holloway,J.L. and Lok,S.
TITLE Secreted protein zacrp4
JOURNAL Patent: WO 0102565-A 4 11-JAN-2001;
ZymoGenetics, Inc. (US)
FEATURES
source Location/Qualifiers
1..987
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="degenerate nucleotide sequence encoding the
polypeptide of SEQ ID NO:2"
1..987
/note="Each N is A, T, G or C"
BASE COUNT 111 a 127 c 204 g 120 t 425 others
ORIGIN
Query Match 4.3%; Score 15; DB 6; Length 987;
Best Local Similarity 100.0%; Pred. No. 4.9e-19;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 337 GCNTGYTGCGCNYTN 351
|||||
Db 37 GCNTGYTGCGCNYTN 51
RESULT 9
AR108644
LOCUS AR108644 1155 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 3 from patent US 6111075.
ACCESSION AR108644
VERSION AR108644.1 GI:12824131
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1155)
AUTHORS Xu,W.-f., Presnell,S.R., Yee,D.P. and Foster,D.C.
TITLE Protease-activated receptor PAR4 (ZCHEMR2)
JOURNAL Patent: US 6111075-A 3 29-AUG-2000;
FEATURES
source Location/Qualifiers
1..1155
/organism="unknown"
BASE COUNT 100 a 143 c 217 g 161 t 534 others
ORIGIN
Query Match 4.3%; Score 15; DB 6; Length 1155;
Best Local Similarity 100.0%; Pred. No. 4.6e-19;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 145 TAYWSNGAYCCNWSN 159
|||||
Db 919 TAYWSNGAYCCNWSN 933
RESULT 10
AX179624
LOCUS AX179624 2019 bp DNA linear PAT 06-AUG-2001
DEFINITION Sequence 3 from Patent WO0146418.
ACCESSION AX179624
VERSION AX179624.1 GI:15132050
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 2019)
AUTHORS Holloway,J.L. and Chandrasekher,Y.A.
TITLE Human slit polypeptide zslit3
JOURNAL Patent: WO 0146418-A 3 28-JUN-2001;
ZymoGenetics, Inc. (US)
FEATURES
source Location/Qualifiers
1..2019
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="degenerate polynucleotide sequence for zslit3"

```
BASE COUNT      231 a      297 c      354 g      230 t      907 others
ORIGIN

Query Match      4.3%; Score 15; DB 6; Length 2019;
Best Local Similarity 100.0%; Pred. No. 3.6e-19;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 310 WSNYTNMGNGTNGGN 324
|||||
Db 1423 WSNYTNMGNGTNGGN 1437

RESULT 11
LOCUS      AR004980      2277 bp      DNA      linear      PAT 04-DEC-1998
DEFINITION      Sequence 2 from patent US 5747317.
ACCESSION      AR004980
VERSION      AR004980.1 GI:3965859
SOURCE      .
ORGANISM      Unknown.
REFERENCE      1 (bases 1 to 2277)
AUTHORS      Cao,Z.
TITLE      Human telomerase RNA interacting protein gene
JOURNAL      Patent: US 5747317-A 2 05-MAY-1998;
FEATURES      Location/Qualifiers
             source
             1..2277
             /organism="unknown"

BASE COUNT      511 a      212 c      395 g      216 t      943 others
ORIGIN

Query Match      4.3%; Score 15; DB 6; Length 2277;
Best Local Similarity 100.0%; Pred. No. 3.4e-19;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 294 YCARTGGAARCARGA 308
|||||
Db 2106 YCARTGGAARCARGA 2120

RESULT 12
LOCUS      I26713      20 bp      DNA      linear      PAT 07-OCT-1996
DEFINITION      Sequence 14 from patent US 5559220.
ACCESSION      I26713
VERSION      I26713.1 GI:1606583
SOURCE      .
ORGANISM      Unknown.
REFERENCE      1 (bases 1 to 20)
AUTHORS      Roessler,P.G. and Ohlrogge,J.B.
TITLE      Gene encoding acetyl-coenzyme A carboxylase
JOURNAL      Patent: US 5559220-A 14 24-SEP-1996;
FEATURES      Location/Qualifiers
             source
             1..20
             /organism="unknown"

BASE COUNT      1 a      5 c      3 g      5 t      6 others
ORIGIN

Query Match      4.0%; Score 14; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 9.1e-16;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 CAYGGAAYGCNGT 35
|||||
Db 14 CAYGGAAYGCNGT 1

RESULT 13
AX350993
LOCUS
```

```
DEFINITION      Sequence 27 from Patent WO0190358.
ACCESSION      AX350993
VERSION      AX350993.1 GI:18616369
KEYWORDS      .
SOURCE      unidentified.
ORGANISM      unidentified.
REFERENCE      1
AUTHORS      Gorman,D.M.
TITLE      Mammalian receptor proteins; related reagents and methods
JOURNAL      Patent: WO 0190358-A 27 29-NOV-2001;
FEATURES      Location/Qualifiers
             source
             1..1026
             /organism="unidentified"
             /db_xref="taxon:32644"
             /note="rodent; surmised Mus musculus"

BASE COUNT      170 a      157 c      156 g      122 t      421 others
ORIGIN

Query Match      4.0%; Score 14; DB 6; Length 1026;
Best Local Similarity 100.0%; Pred. No. 1.6e-16;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 257 AYYTNCCNGARGAR 270
|||||
Db 515 AYYTNCCNGARGAR 528

RESULT 14
LOCUS      AX172419      1134 bp      DNA      linear      PAT 03-JUL-2001
DEFINITION      Sequence 3 from Patent WO0144479.
ACCESSION      AX172419
VERSION      AX172419.1 GI:14597534
KEYWORDS      .
SOURCE      synthetic construct.
ORGANISM      synthetic construct.
REFERENCE      1 (bases 1 to 1134)
AUTHORS      Conklin,D.C., Yamamoto,G., Cooper,E. and Jaspers,S.R.
TITLE      Galactosyltransferase homolog, znssp8
JOURNAL      Patent: WO 0144479-A 3 21-JUN-2001;
FEATURES      Location/Qualifiers
             source
             1..1134
             /organism="synthetic construct"
             /db_xref="taxon:32630"
             /note="degenerate sequence"

BASE COUNT      230 a      102 c      169 g      174 t      459 others
ORIGIN

Query Match      4.0%; Score 14; DB 6; Length 1134;
Best Local Similarity 100.0%; Pred. No. 1.6e-16;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 326 ARGNTAYGARGCN 339
|||||
Db 800 ARGNTAYGARGCN 813

RESULT 15
AX167244
LOCUS      AX167244      1584 bp      DNA      linear      PAT 03-JUL-2001
DEFINITION      Sequence 3 from Patent WO0144281.
ACCESSION      AX167244
VERSION      AX167244.1 GI:14596712
KEYWORDS      .
SOURCE      synthetic construct.
ORGANISM      synthetic construct.
REFERENCE      1 (bases 1 to 1584)
AUTHORS      Holloway,J.L.
```

TITLE Human secretin-like g-protein coupled receptor
 JOURNAL Patent: WO 0144281-A 3 21-JUN-2001;
 ZymoGenetics, Inc. (US)
 FEATURES Location/Qualifiers
 source
 1..1584
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="This degenerate nucleotide sequence encodes the
 amino acid sequence of SEQ ID NO:2."
 BASE COUNT 210 a 157 c 245 g 257 t 715 others
 ORIGIN
 Query Match 4.0%; Score 14; DB 6; Length 1584;
 Best Local Similarity 100.0%; Pred. No. 1.4e-16;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 56 GNCNMGNGARYTN 69
 |
 Db 374 GNCNMGNGARYTN 387

Search completed: November 7, 2002, 04:21:28
 Job time : 2775 secs

Barrett
82 09 649 66 28
USD-849-050

THIS PAGE BLANK (USPTO)

KW bone marrow transplants; proliferation-restricted cells;
 XX ambiguity-maximised; ds.

OS Homo sapiens.
 XX Synthetic.

FH Key Location/Qualifiers
 FT CDS 1..2277
 FT /*tag= a
 FT /product= telomerase protein p105

XX WO9801543-A1.

XX 15-JAN-1998.

PD 08-JUL-1997; 97WO-US12297.

PF 08-JUL-1996; 96US-0676967.

PR (TULA-) TULARIK INC.

XX Cao Z;

XX WPI; 1998-101044/09.

DR P-PSDB; AAW41927.

XX New nucleic acid encoding human telomerase protein p105 or its
 PT fragments - used for therapeutic modulation of telomerase activity
 PT and for screening for potential modulators of telomerase-target
 PT binding

XX Disclosure; Pages 20-21; 32pp; English.

XX The sequence is that of an ambiguity-maximised human telomerase
 CC protein coding sequence. The sequence, or specific fragments
 CC of it, can be used to modulate expression of a telomerase
 CC transcript (by hybridising to it intracellularly), e.g.
 CC for treatment or prevention of cancer, restenosis, inflammation,
 CC myocardial infarction, glomerulonephritis, transplant rejection and
 CC infections (e.g. with human immunodeficiency virus). It can be used to
 CC express recombinant telomerase protein which can be used to screen
 CC for agents, e.g. antibodies, that modulate binding of human telomerase
 CC to its binding target. Those that inhibit telomerase activity can be
 CC used to treat the conditions listed above, while those that are
 CC agonists can be used to extend the life of proliferation-restricted
 CC cells, especially normal somatic cells, e.g. in cases of
 CC hypersensitivity or atrophy, also to improve production of
 CC recombinant proteins by maximising cell density and survival and
 CC expansion of precursor cells being used for bone marrow transplants.
 CC They may also be used for diagnosis. Other uses of telomerase proteins
 CC are isolation, enrichment and concentration of telomerase RNA or
 CC proteins; as immunogens; in therapy; as reagent where nascent
 CC oligonucleotides of known structure are needed (e.g. for tagging
 CC native nucleic acid molecules) and for regulating cell growth/density
 CC tolerance. The agents and the telomerase proteins should be very
 CC specific, e.g. they are selective for cancer cells without harming
 CC somatic cells.

XX Sequence 2277 BP; 513 A; 212 C; 395 G; 217 T; 940 other;

Query Match 4.3%; Score 15; DB 19; Length 2277;
 Best Local Similarity 100.0%; Pred. No. 0.003;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 294 YCARTGGAARCARGA 308

Db 2106 YCARTGGAARCARGA 2120

RESULT 15

AAV05370

ID AAV05370 standard; cDNA; 2277 BP.

XX

AC AAV05370;

XX 06-JUL-1998 (first entry)

DE Human telomerase p105 subunit synthetic gene.

XX Telomerase; p105; human; cell replication; cancer; restenosis;
 KW multiple sclerosis; inflammation; rheumatoid arthritis;
 KW myocardial infarction; glomerulonephritis; transplant rejection;
 KW infection; therapy; ds.

XX Homo sapiens.

OS Synthetic.

XX WO9801542-A1.

XX 15-JAN-1998.

XX 08-JUL-1997; 97WO-US12296.

XX 08-JUL-1996; 96US-0676974.

XX (REGC) UNIV CALIFORNIA.

XX Collins K;

XX WPI; 1998-101043/09.

XX New nucleic acid encoding human telomerase proteins or their
 PT fragments - useful for therapeutic modulation of telomerase activity
 PT and for screening for potential modulators of telomerase-target
 PT binding

XX Disclosure; Page 19-20; 32pp; English.

XX This polynucleotide comprises a synthetic, ambiguity-maximised
 CC DNA coding for the p105 subunit (see AAW46593) of human telomerase.
 CC It is based on an isolated cDNA clone (see AAV05369) for p105 and
 CC encompasses all possible nucleic acids encoding the full-length
 CC protein. The invention provides methods relating to human
 CC telomerase and related nucleic acids, including the subunit
 CC proteins p140, p105, p48 and p43. The proteins may be produced
 CC recombinantly from transformed host cells or purified from human
 CC cells. Also included are human telomerase RNA (see AAV05373) and
 CC functional derivatives (see AAV05374 and AAV16092-93), as well as p105
 CC synthetic DNA sequences (AAV05370-72). The invention also provides
 CC isolated telomerase hybridisation probes and primers capable of
 CC specifically hybridising with the telomerase gene, telomerase-
 CC specific binding agents such as specific antibodies, and methods of
 CC making and using the subject compositions in diagnosis (e.g.
 CC genetic hybridisation screens for telomerase transcripts), therapy
 CC (e.g. gene therapy to modulate telomerase gene expression) and in
 CC the biopharmaceutical industry (e.g. reagents for screening
 CC chemical libraries for lead agents). Modulation of telomerase
 CC expression can be used for the treatment or prevention of cancer,
 CC restenosis, inflammation, myocardial infarction, glomerulonephritis,
 CC transplant rejection or infections (e.g. with HIV).

XX Sequence 2277 BP; 511 A; 212 C; 395 G; 216 T; 943 other;

Query Match 4.3%; Score 15; DB 19; Length 2277;
 Best Local Similarity 100.0%; Pred. No. 0.003;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 294 YCARTGGAARCARGA 308

Db 2106 YCARTGGAARCARGA 2120

Search completed: November 7, 2002, 03:35:06

Job time : 307 secs

KW platelet; proliferation; differentiation; mediation;
 KW inflammatory process; vascular injury; chemotaxis; mitogenesis;
 KW growth factor; production; degenerate; ds.
 XX Synthetic.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FH CDS 1..1155
 FT /*tag= a
 FT /product= "Human protease-activated receptor PAR4"
 FT /note= "No stop codon given in the specification"
 XX
 XX WO9950415-A2.
 XX 07-OCT-1999.
 XX 31-MAR-1999; 99WO-US07100.
 XX 01-APR-1998; 98US-0053866.
 XX (ZYMO) ZYMOGENETICS INC.
 PA (UNITW) UNIV WASHINGTON.
 XX Xu W, Presnell SR, Yee DP, Foster DC;
 XX WPI: 1999-633640/54.
 DR P-PSDB; AAY50135.
 XX Novel protease activated receptor 4, useful for screening for
 PT (antagonists for promoting the proliferation and/or differentiation of
 PT platelets and in mediating inflammatory events -
 XX Disclosure; Page 81-82; 85pp; English.
 XX This sequence represents degenerate DNA encoding human protease-activated
 CC receptor PAR4 (also referred to as ZCHEMR2). Protease-activated receptors
 CC (PARs) are a subfamily of G protein coupled receptors which are capable
 CC of mediating cellular signalling in response to proteases (e.g.,
 CC thrombin). They are characterised by a tethered peptide ligand at the
 CC extracellular N-terminus that is generated by proteolysis. PAR4 is
 CC activated by thrombin or trypsin cleavage at Arg47/Gly48, which generates
 CC a new N-terminus corresponding to the tethered ligand (a hexapeptide).
 CC Agonists of PAR4 are useful for upregulating cellular or physiological
 CC responses whereas antagonists are used to downregulate these
 CC activities. The PAR4 protein is further useful for dissecting the
 CC effects of thrombin or other activating proteases in the clotting
 CC pathway from the effects of these proteases at the cellular level.
 CC Agonists are specifically useful in promoting the proliferation
 CC and/or differentiation of platelets, in promoting inflammatory events,
 CC responses to vascular injury, chemotaxis or mitogenesis, and in
 CC producing growth factors. Antagonists are useful as research reagents
 CC for characterising sites of ligand-receptor interaction.
 XX Sequence 1155 BP; 100 A; 143 C; 217 G; 161 T; 534 other;
 SQ
 Query Match 4.3%; Score 15; DB 20; Length 1155;
 Best Local Similarity 100.0%; Pred. No. 0.0038;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 145 TAYWSNGAYCCNWSN 159
 DB 919 TAYWSNGAYCCNWSN 933
 RESULT 13
 AAH28471
 ID AAH28471 standard; DNA; 2019 BP.
 XX
 AC AAH28471;
 XX
 XX 17-SEP-2001 (first entry)
 XX

DE Degenerate nucleotide sequence of human slit polypeptide Zslit3.
 XX
 KW Slit protein; Zslit3; neurite growth; cellular proliferation;
 KW immune response; stroke; brain damage; paralysis; Huntington's disease;
 KW neurodegenerative disease; amyotrophic lateral sclerosis;
 KW Alzheimer's disease; Parkinson's disease; peripheral neuropathy;
 KW demyelinating disease; multiple sclerosis; lung organogenesis;
 KW pulmonary disease; respiration; sclerosis; cystic fibrosis; asthma;
 KW immunosuppression; autoimmune disease; insulin dependent diabetes;
 KW rheumatoid arthritis; ss.
 XX Homo sapiens.
 XX WO200146418-A1.
 XX 28-JUN-2001.
 XX 14-DEC-2000; 2000WO-US34230.
 XX 21-DEC-1999; 99US-0469847.
 XX (ZYMO) ZYMOGENETICS INC.
 PA Holloway JL, Chandrasekher YA;
 XX WPI: 2001-441677/47.
 XX Novel human slit polypeptide, ZSLIT3, useful for treating and
 PT diagnosing cystic fibrosis, insulin dependent diabetes and multiple
 PT sclerosis -
 XX Claim 3; Page 122; 125pp; English.
 XX The present sequence encodes a human slit protein polypeptide,
 CC designated Zslit3. Zslit is a neurite growth and development modulator,
 CC and an cellular proliferation and differentiation and immune response
 CC modulator. Zslit3 polypeptides and polynucleotides are useful for
 CC regenerating and directing neurite outgrowths following strokes, brain
 CC damage caused by head injuries, paralysis caused by spinal injuries,
 CC and for treating neurodegenerative diseases such as amyotrophic lateral
 CC sclerosis, Alzheimer's disease, Huntington's disease, Parkinson's
 CC disease and peripheral neuropathies, or demyelinating diseases
 CC e.g. multiple sclerosis. They are useful for lung organogenesis and
 CC repair, and thus useful for diagnosing and treating pulmonary diseases
 CC such as respiration and circulation, cystic fibrosis and asthma. They
 CC also act as a mediator of immunosuppression, and thus are useful for
 CC diagnosing and treating autoimmune diseases such as insulin dependent
 CC diabetes and rheumatoid arthritis.
 XX Sequence 2019 BP; 231 A; 297 C; 354 G; 230 T; 907 other;
 SQ
 Query Match 4.3%; Score 15; DB 22; Length 2019;
 Best Local Similarity 100.0%; Pred. No. 0.0031;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 310 WSNYTNMGNGTNGGN 324
 DB 1423 WSNYTNMGNGTNGGN 1437
 RESULT 14
 AAV13834
 ID AAV13834 standard; cDNA; 2277 BP.
 XX
 AC AAV13834;
 XX
 XX 21-JUL-1998 (first entry)
 XX Homo sapiens ambiguity-maximised telomerase protein p105 gene.
 DE telomerase; p105; treatment; prevention; cancer; restenosis;
 KW inflammation; myocardial infarction; glomerulonephritis; transplant;
 KW rejection; infection; HIV; human immunodeficiency virus;

Qy 257 AYTNCNGARG 271
Db 482 AYTNCNGARG 496
|||||

RESULT 10

ABK15336
ID ABK15336 standard; DNA: 729 BP.

XX AC ABK15336;

DT 08-MAY-2002 (first entry)

XX Human helical protein zalpha51 degenerate coding sequence #2.

KW Human: helical protein; zalpha51; gene; cytokine; muscular;
DE neuroprotective; gene therapy; vaccine; protein therapy;
KW neuromuscular disorder; locomotion disorder; ds.

XX OS Homo sapiens.

PN US2002009775-A1.

XX PD 24-JAN-2002.

XX PF 16-MAR-2001; 2001US-0810052.

XX PR 17-MAR-2000; 2000US-190410P.

XX PR 25-APR-2000; 2000US-199443P.

XX PA (CONK/) CONKLIN D C.

PA (PRES/) PRESNELL S R.

XX PI Conklin DC, Presnell SR;

XX DR WPI; 2002-171143/22.

XX Helical protein zalpha51 and an isolated Glu-Glu affinity tag
PT polypeptide, useful in the prevention, diagnosis and treatment of
PT neuromuscular disorders, e.g. locomotion disorders -

PS Claim 28; Page 31; 4lpp; English.

XX The present invention relates to a new helical protein zalpha51
CC (comprising a 4 helix bundle) and an isolated Glu-Glu affinity tag
CC polypeptide, and the nucleic acids encoding them, as defined in the
CC specification. The nucleic acids and proteins may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate zalpha51 expression. The nucleic acids and their
CC complements may also be used as DNA probes in diagnostic assays to
CC detect and quantitate the presence of similar nucleic acids in samples,
CC and therefore which patients may be in need of restorative therapy. The
CC zalpha51 polypeptide may also be used as antigens in the production of
CC antibodies against zalpha51 and in assays to identify modulators of
CC zalpha51 expression and activity. The anti-zalpha51 antibodies and
CC antagonists may also be used to down regulate expression and activity.
CC The anti-zalpha51 antibodies may also be used as diagnostic agents for
CC detecting the presence of zalpha51 polypeptide in samples. Disorders that
CC may be prevented, diagnosed and/or treated by the above methods include,
CC for example, neuromuscular disorders, especially a locomotion disorder.
CC The present nucleic acid sequence represents all nucleic acids that
CC encode zalpha51 #2.

XX SQ Sequence 729 BP; 78 A; 85 C; 133 G; 96 T; 337 other;

Query Match 4.3%; Score 15; DB 24; Length 729;
Best Local Similarity 100.0%; Pred. No. 0.0045;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 257 AYTNCNGARG 271

Db 482 AYTNCNGARG 496

|||||

RESULT 11

AAF28673
ID AAF28673 standard; CDNA; 987 BP.

XX AC AAF28673;

DT 05-APR-2001 (first entry)

XX Degenerate human zacr4 coding sequence.

XX Human; zacr4; complement factor C1q domain; chromosome 11q11;
KW energy balance; cellular metabolic reaction; autocrine factor;
DE development; cell proliferation; differentiation; cell survival; ss.

XX OS Homo sapiens.

PN W0200102565-A2.

XX PD 11-JAN-2001.

XX PF 28-JUN-2000; 2000WO-US17692.

XX PR 01-JUL-1999; 99US-0346502.

XX PA (ZYMO) ZYMOGENETICS INC.

XX PI Holloway JL, Lok S;

XX DR WPI; 2001-138140/14.

XX DR P-PSDB; AAB61606.

XX PT Novel secreted protein ZACRP4 polypeptides having tandem C1q globular
PT domains, useful for studying cell-cell communication and regulation of
PT cellular processes -

XX PS Disclosure; Page 80; 82pp; English.

XX The present sequence is a degenerate coding sequence for human ZACRP4
CC protein. ZACRP4 protein has two complement factor C1q domains. The ZACRP4
CC gene is located on human chromosome 11q11. The ZACRP4 coding sequence and
CC protein have a number of uses described in the specification, including,
CC modulation of energy balance and cellular metabolic reactions in
CC mammals. In addition, ZACRP4 protein is useful as an autocrine factor,
CC particularly during development, in mediating the processes of an
CC organism, in regulating cellular processes such as cell proliferation
CC and/or differentiation, cell survival and energy balance.

XX SQ Sequence 987 BP; 111 A; 127 C; 204 G; 120 T; 425 other;

Query Match 4.3%; Score 15; DB 22; Length 987;

Best Local Similarity 100.0%; Pred. No. 0.0041;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 337 GCNTGTYGGGCGNYTN 351

Db 37 GCNTGTYGGGCGNYTN 51

|||||

RESULT 12

AAZ32748
ID AAZ32748 standard; DNA; 1155 BP.

XX AC AAZ32748;

DT 31-JAN-2000 (first entry)

XX Human protease-activated receptor PAR4 degenerate DNA.

KW Protease-activated receptor; PAR4; ZCHEMR2; G protein coupled;
KW cellular signalling; protease; tethered ligand; N-terminal;
KW proteolysis; thrombin; trypsin; cleavage; hexapeptide; agonist;
KW antagonist; cellular response; physiological response; clotting pathway;

CC degenerate DNA #1 encompassing all DNAs encoding human zalpha51.
XX Sequence 696 BP; 74 A; 82 C; 121 G; 93 T; 326 other;
SQ

Query Match 4.3%; Score 15; DB 22; Length 696;
Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 AYYTNCNGARGARG 271
DB 449 AYYTNCNGARGARG 463
|||||

RESULT 8
ABK15334
ID ABK15334 standard; DNA; 696 BP.
XX
AC ABK15334;
XX

DT 08-MAY-2002 (first entry)
XX
DE Human helical protein zalpha51 degenerate coding sequence #1.
XX
KW Human; helical protein; zalpha51; gene; cytokine; muscular;
KW neuroprotective; gene therapy; vaccine; protein therapy;
KW neuromuscular disorder; locomotion disorder; ds.
XX
OS Homo sapiens.
XX

PN US2002009775-A1.
XX
PD 24-JAN-2002.
XX

PF 16-MAR-2001; 2001US-0810052.
XX
PR 17-MAR-2000; 2000US-190410P.
PR 25-APR-2000; 2000US-199443P.
XX

PA (CONK/) CONKLIN D.C.
PA (PRES/) PRESNELL S.R.
XX

PI Conklin DC, Presnell SR;
XX

DR WPI; 2002-171143/22.
XX

PT Helical protein zalpha51 and an isolated Glu-Glu affinity tag
PT polypeptide, useful in the prevention, diagnosis and treatment of
PT neuromuscular disorders, e.g. locomotion disorders -
XX

PS Disclosure; Page 28-29; 41pp; English.
XX

CC The present invention relates to a new helical protein zalpha51
CC (comprising a 4 helix bundle) and an isolated Glu-Glu affinity tag
CC polypeptide, and the nucleic acids encoding them, as defined in the
CC specification. The nucleic acids and proteins may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate zalpha51 expression. The nucleic acids and their
CC complements may also be used as DNA probes in diagnostic assays to
CC detect and quantitate the presence of similar nucleic acids in samples,
CC and therefore which patients may be in need of restorative therapy. The
CC zalpha51 polypeptide may also be used as antigens in the production of
CC antibodies against zalpha51 and in assays to identify modulators of
CC zalpha51 expression and activity. The anti-zalpha51 antibodies and
CC antagonists may also be used to down regulate expression and activity.
CC The anti-zalpha51 antibodies may also be used as diagnostic agents for
CC detecting the presence of zalpha51 polypeptide in samples. Disorders that
CC may be prevented, diagnosed and/or treated by the above methods include,
CC for example, neuromuscular disorders, especially a locomotion disorder.
CC The present nucleic acid sequence represents all nucleic acids that
CC encode zalpha51 #1.
XX

SQ Sequence 696 BP; 74 A; 82 C; 121 G; 93 T; 326 other;

Query Match 4.3%; Score 15; DB 24; Length 696;
Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 AYYTNCNGARGARG 271
DB 449 AYYTNCNGARGARG 463
|||||

RESULT 9
AAS13926
ID AAS13926 standard; DNA; 729 BP.
XX
AC AAS13926;
XX

DT 19-DEC-2001 (first entry)
XX

DE Degenerate DNA #2 encompassing all DNAs encoding human zalpha51.
XX
KW Human; zalpha51; antirheumatic; antiarthritic; neuroprotective;
KW antidiabetic; antiinflammatory; dermatological; immunosuppressive;
KW antiasthmatic; antibacterial; cytostatic; neuromuscular disorder;
KW locomotion disorder; cancer; tumour imaging; immune disorder;
KW graft rejection; graft-versus-host disease; ds.
XX
OS Synthetic.
XX

PN WO200170986-A2.
XX

PD 27-SEP-2001.
XX

PF 16-MAR-2001; 2001WO-US08493.
XX

PR 17-MAR-2000; 2000US-0527843.
PR 25-APR-2000; 2000US-0558459.
XX

PA (ZYMO) ZYMOGENETICS INC.
XX

PI Conklin DC, Presnell SR;
XX

DR WPI; 2001-639129/73.
XX

PT Novel cytokine, zalpha51, useful as diagnostic for liver,
PT neuroblastoma, brain and other cancers, and for treating rheumatoid
PT arthritis, multiple sclerosis, myasthenia gravis, systemic lupus
PT erythematosus, diabetes -
XX

PS Claim 28; p89; 89pp; English.
XX

CC The invention relates to a novel isolated zalpha 51 polypeptide (a
CC cytokine) (I). The nucleic acid (II) encoding zalpha51 is useful as
CC a probe for detecting the presence of RNA encoding (I) in a biological
CC sample. (II) is also useful as a diagnostic reagent for detecting a
CC genetic abnormality in a patient. The antibody to (I) is useful for
CC detecting the presence of (I) in a biological sample taken from a
CC mammal with a neuromuscular disorder or locomotion disorder. (I) is
CC useful for raising antibodies, including monoclonal antibodies
CC that specifically bind to (I). The zalpha51 antagonist is useful for
CC treating cancer, and also for modulating the immune system, reducing
CC graft rejection, preventing graft-versus-host disease, boosting
CC immunity to infectious diseases, treating immuno-compromised patients
CC or improving vaccines. (I), (II), and the antibody to (I) are useful
CC for diagnosing and treating disorders associated with abnormal cell
CC proliferation. Labeled zalpha51 polypeptides may also be used for
CC imaging tumours. The present sequence represents the coding sequence of
CC degenerate DNA #2 encompassing all DNAs encoding human zalpha51.
XX

SQ Sequence 729 BP; 78 A; 85 C; 133 G; 96 T; 337 other;

Query Match 4.3%; Score 15; DB 22; Length 729;
Best Local Similarity 100.0%; Pred. No. 0.0045;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX SQ Sequence 2379 BP; 270 A; 286 C; 502 G; 272 T; 1049 other;
Query Match 4.6%; Score 16; DB 22; Length 2379;
Best Local Similarity 100.0%; Pred. No. 0.00043;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 123 NWSNGNWSNTGYMGN 138
Db 1512 NWSNGNWSNTGYMGN 1527
|||||

RESULT 6
AAC91387
ID AAC91387 standard; DNA; 426 BP.
XX AC
XX AAC91387;
XX
XX
XX 19-MAR-2001 (first entry)
XX
XX Human zalpha31 degenerate coding sequence.
XX
XX Human: zalpha31; alpha helical protein-31;
KW Cytostatic; antiinflammatory; antiarthritic; antirheumatoid;
KW immunosuppressive; antidiabetic; antibacterial; osteopathic;
KW antithrombotic; antipsoriatic; dermatological; anti-HIV;
KW gene therapy; osteoporosis; Paget's disease; hyperparathyroidism;
KW acute pancreatitis; gastrointestinal disorder; inflammation;
KW heart disorder; viral infection; ds.
XX
XX Homo sapiens.
OS Synthetic.
OS
XX WO200073458-A1.
XX
XX
XX PD 07-DEC-2000.
XX
XX PF 26-MAY-2000; 2000WO-US14795.
XX
XX PR 28-MAY-1999; 99US-0136485.
XX
XX PA (ZYMO) ZYMOGENETICS INC.
XX
XX PI Conklin DC;
XX
XX WPI; 2001-070967/08.
XX P-PSDB; AAB50867.
XX
XX PT Novel 4-helix bundle cytokine, Zalpha31, useful for regulating the
PT function of immune system and for treating thyroid, adrenal, lymphoid,
PT inflammatory, pancreatic, blood or bone disorders .
XX
XX Claim 3; Page 100-101; 111pp; English.

The present sequence is given in a specification relating to polynucleotide and polypeptide molecules for mammalian secreted alpha helical protein-31 (zalpha31). The polypeptides are a novel four-helix bundle cytokine and may be used to regulate the functioning of the immune system. The polypeptides are used to identify and isolate receptors involved in spermatogenesis, steroidogenesis, testicular differentiation and regulatory control of the hypothalamic-pituitary gonadal axis, thyroid, heart and adrenal function. They are useful for treating disorders of the reproductive system, thyroid, adrenal, heart and immunological systems. Zalpha31 polypeptides, or antagonists are useful in the treatment of e.g. osteoporosis, Paget's disease, and hyperparathyroidism, acute pancreatitis and gastrointestinal disorders and as analgesics, especially for bone pain. They can be used to treat atherosclerosis, pelvic inflammatory disease, (PID), psoriasis, arthritis, eczema, scleroderma and other inflammatory diseases. The polypeptides, nucleic acids and/or antibodies can be used to treat heart disorders. They can limit infarct size following a heart attack, aid in recovery after heart transplantation, treat complications related to poor circulation e.g. diabetic foot ulcers, treat stroke. Improve

CC cardiac function, induce skeletal muscle neogenesis and/or hyperplasia,
CC be used in kidney regeneration and/or for treating systemic and
CC pulmonary hypertension. Zalpha31 can be used for treating viral
CC leukaemias, acquired immunodeficiency syndrome (AIDS) or gastrointestinal
CC viral infections.
XX
SQ Sequence 426 BP; 101 A; 41 C; 70 G; 43 T; 171 other;
Query Match 4.3%; Score 15; DB 22; Length 426;
Best Local Similarity 100.0%; Pred. No. 0.0055;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 263 CNGARGARGCNAARG 277
Db 77 CNGARGARGCNAARG 91
|||||

RESULT 7
AAS13924
ID AAS13924 standard; DNA; 696 BP.
XX AC
XX AAS13924;
XX
XX 19-DEC-2001 (first entry)
XX
XX Degenerate DNA #1 encompassing DNAs encoding human Zalpha51.
XX
XX Human: Zalpha51; antirheumatic; antiarthritic; neuroprotective;
KW antidiabetic; antiinflammatory; dermatological; immunosuppressive;
KW antithrombotic; antibacterial; cytostatic; neuromuscular disorder;
KW locomotion disorder; cancer; tumour imaging; immune disorder;
KW graft rejection; graft-versus-host disease; ds.
XX
XX Synthetic.
OS
XX WO200170986-A2.
XX
XX PD 27-SEP-2001.
XX
XX PF 16-MAR-2001; 2001WO-US08493.
XX
XX PR 17-MAR-2000; 2000US-0527843.
XX 25-APR-2000; 2000US-0558459.
XX
XX PA (ZYMO) ZYMOGENETICS INC.
XX
XX PI Conklin DC, Presnell SR;
XX
XX WPI; 2001-639129/73.
XX
XX Novel cytokine, Zalpha51, useful as diagnostic for liver,
XX neuroblastoma, brain and other cancers, and for treating rheumatoid
XX arthritis, multiple sclerosis, myasthenia gravis, systemic lupus
XX erythematosus, diabetes
XX
XX Disclosure: p85; 89pp; English.

The invention relates to a novel isolated Zalpha 51 polypeptide (a cytokine) (I). The nucleic acid (II) encoding Zalpha51 is useful as a probe for detecting the presence of RNA encoding (I) in a biological sample. (II) is also useful as a diagnostic reagent for detecting a genetic abnormality in a patient. The antibody to (I) is useful for detecting the presence of (I) in a biological sample taken from a mammal with a neuromuscular disorder or locomotion disorder. (I) is useful for raising antibodies, including monoclonal antibodies that specifically bind to (I). The Zalpha51 antagonist is useful for treating cancer, and also for modulating the immune system, reducing graft rejection, preventing graft-versus-host disease, boosting immunity to infectious diseases, treating immuno-compromised patients or improving vaccines. (I), (II), and the antibody to (I) are useful for diagnosing and treating disorders associated with abnormal cell proliferation. Labeled Zalpha51 polypeptides may also be used for imaging tumours. The present sequence represents the coding sequence of

CC OX2RH1, OX2RH2, OX2RH3, OX2RH4, or OX2RH1.2 proteins. Agonists and
CC antagonists of the OX2RH sequences can be used to modulate physiology
CC or development of a cell, particularly for enhancing myeloid function or
CC enhancing immunity. The sequences can be used to identify non-OX2 ligands
CC for an OX2R. The polypeptides and polynucleotides can be used to treat
CC inflammatory, leukoproliferative, neurodegenerative or post-traumatic
CC conditions, including atherosclerosis, multiple sclerosis, ischaemia,
CC neurodegeneration, rheumatoid arthritis, and autoimmunity. The present
CC sequence represents the reverse translated degenerate polynucleotide
CC sequence of human OX2R homologue 1 (OX2RH1).

XX SQ Sequence 885 BP; 182 A; 98 C; 133 G; 114 T; 358 other;

Query Match 4.6%; Score 16; DB 22; Length 885;

Best Local Similarity 100.0%; Pred. No. 0.00061; Length 885;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 229 AAYGTNWSNACNGTNA 244

|||||

Db 697 AAYGTNWSNACNGTNA 712

RESULT 4

AAC84178

ID AAC84178 standard; DNA: 1044 BP.

AC AAC84178;

XX 19-MAR-2001 (first entry)

DE Human OX2RH1.2 degenerate nucleotide sequence.

XX OX2R protein: OX2RH1; OX2RH2; OX2RH3; OX2RH4; OX2RH1.2; ischaemia;
KW antiinflammatory; cytostatic; neuroprotective; nootropic; human;
KW antiarteriosclerotic; vasotropic; immunosuppressive; antirheumatic;
XX antiarthritic; gene therapy; ds.

OS Homo sapiens.

XX W0200070045-A1.

PN 23-NOV-2000.

XX 11-MAY-2000; 2000WO-US12998.

XX 13-MAY-1999; 99GB-0011123.

PR 03-NOV-1999; 99GB-0025989.

XX (MEDI-) MEDICAL RES COUNCIL.

PA (SCHE) SCHERING CORP.

XX Barclay AN, Brown MH, Gorman DM, Lanier LL, Wright GJ;

PI Cherwinski H, Phillips JH, Hoek RM, Sedgwick JD;

XX WPI; 2001-016233/02.

DR P-PSDB; AAB48016.

XX Mammalian OX2R proteins and DNA sequences useful for modulating the
PT physiology and development of a cell -

XX Disclosure; Page 36; 142pp; English.

XX The invention relates to rodent or primate OX2R proteins, especially
CC OX2RH1, OX2RH2, OX2RH3, OX2RH4, or OX2RH1.2 proteins. Agonists and
CC antagonists of the OX2RH sequences can be used to modulate physiology
CC or development of a cell, particularly for enhancing myeloid function or
CC enhancing immunity. The sequences can be used to identify non-OX2 ligands
CC for an OX2R. The polypeptides and polynucleotides can be used to treat
CC inflammatory, leukoproliferative, neurodegenerative or post-traumatic
CC conditions, including atherosclerosis, multiple sclerosis, ischaemia,
CC neurodegeneration, rheumatoid arthritis, and autoimmunity. The present
CC sequence represents the reverse translated degenerate polynucleotide
CC sequence of human OX2R homologue 1.2 (OX2RH1.2).

XX SQ Sequence 1044 BP; 220 A; 117 C; 154 G; 131 T; 422 other;

Query Match 4.6%; Score 16; DB 22; Length 1044;

Best Local Similarity 100.0%; Pred. No. 0.00058;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 229 AAYGTNWSNACNGTNA 244

|||||

Db 697 AAYGTNWSNACNGTNA 712

RESULT 5

AAH74766

ID AAH74766 standard; DNA: 2379 BP.

XX AAH74766;

XX 29-OCT-2001 (first entry)

DE Degenerate nucleotide sequence of a human zdint5 polypeptide.

XX Human; zdint5; anti-angiogenic; intestinal polypeptide; wound healing;
KW extracellular matrix interaction; tumour suppression; gamete maturation;
KW immunologic recognition; gastrointestinal irradiation; chemotherapy;
KW proteolysis; apoptosis; angiogenesis; infection; cell adhesion;
KW cell fusion; cell signalling; tumour; Crohn's disease; melanoma; trauma;
KW inflammatory bowel disease; food poisoning; degenerative disease;
KW inflammation; fertility; gamete maturation; epithelial disorder; ss.

XX Synthetic.

OS Homo sapiens.

XX W0200159112-A1.

XX 16-AUG-2001.

XX 09-FEB-2001; 2001WO-US04198.

XX 10-FEB-2000; 2000US-0501806.

PR (ZYMO) ZYMOGENETICS INC.

XX Holloway JL, Sheppard PO;

XX WPI; 2001-522477/57.

XX New anti-angiogenic intestinal polypeptides, zdint5 polypeptides, which
PT are members of disintegrin proteases, for modulating extracellular
PT matrix interaction, tumour suppression and wound healing -

XX Disclosure; Page 88-89; 92pp; English.

XX The present sequence encodes a human zdint5 polypeptide. The zdint5
CC polypeptide is an anti-angiogenic intestinal polypeptide. Zdint5 is
CC used for modulating extracellular matrix interactions. Zdint5
CC polypeptide is useful as a tool for identifying new family members
CC of polypeptides. Zdint5 polynucleotides are useful as probes or primers
CC to clone 5' non-coding regions of zdint5 gene. Zdint5 polypeptides are
CC used for tumour suppression, gamete maturation, immunologic recognition,
CC and growth and differentiation either working in isolation or in
CC conjunction with other molecules in colon, small intestine, fetal lung,
CC testis and B-cells. Zdint5 polypeptides are also useful for promoting
CC wound healing, in the treatment of disorders associated with recovery
CC after gastrointestinal irradiation, chemotherapy or antibody use,
CC as anti-infectives, and extracellular matrix repair and remodeling. The
CC polypeptides are also useful for modulating proteolysis, apoptosis,
CC angiogenesis, infection, cell adhesion, cell fusion and signalling.
CC The polypeptides are also useful for treating tumour formation, Crohn's
CC disease, inflammatory bowel disease, food poisoning, melanoma,
CC degenerative diseases, disorders related to immunity, inflammation,
CC fertility, gamete maturation, immunology, trauma and epithelial
CC disorders.

PT neurodegenerative disorders or cancers
 XX
 PS Example 1; Fig 2; 87pp; English.
 XX
 CC Primer ROS1 is based on an amino acid sequence of the receptor
 CC tyrosine kinase (RTK) ROS. A series of primers (see AAV42003-11)
 CC based on unique motifs of LMR1_r (see AAW48841) or on unrelated
 CC RTKs TEK and ROS were used in the PCR amplification of cDNA from
 CC rat, human and mouse sources. cDNA clones (see AAV32448-56) coding
 CC for novel RTKs designated LMR1, LMR2 and LMR3 (see AAW48841-49)
 CC were obtained. LMRs belong to a unique family of receptors and
 CC are expressed in neuronal tissues and tumour cells, making them
 CC targets for the treatment of neurodegenerative disorders and
 CC cancer.
 XX
 XX Sequence 23 BP; 3 A; 1 C; 8 G; 4 T; 7 other;
 SQ
 Query Match 4.6%; Score 16; DB 19; Length 23;
 Best Local Similarity 100.0%; Pred. No. 0.0023;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 322 GGGGARGTNTAYGARG 337
 Db 7 GGGGARGTNTAYGARG 22
 RESULT 2
 AA240496
 ID AA240496 standard; DNA; 23 BP.
 XX
 AC AA240496;
 XX
 XX
 DT 18-FEB-2000 (first entry)
 XX
 DE Degenerate primer ROS1 for STE20 related protein kinase genes.
 XX
 KW Antirheumatic; antiarthritic; antiinflammatory; antiallergic; osteopathic;
 KW antiproliferative; antiarteriosclerotic; antiasthmatic; immunosuppressive;
 KW neuroprotective; cardiac; cerebroprotective; cytoskeletal; antidiabetic;
 KW vulnery; STE20; protein kinase; STK2; STK3; STK4; STK5; STK6; STK7;
 KW ZC1; ZC2; ZC3; ZC4; KHS2; SULU1; SULU3; GEK2; PAK4; PAK5; antagonist;
 KW antibody; gene therapy; rheumatoid arthritis; artherosclerosis; asthma;
 KW inflammatory bowel disease; Crohn's disease; osteoarthritis; psoriasis;
 KW rhinitis; autoimmunity; organ transplantation; multiple sclerosis;
 KW myocardial infarction; cardiovascular disease; stroke; renal failure;
 KW oxidative stress-related neurodegenerative disorder; Parkinson's disease;
 KW amyotrophic lateral sclerosis; Leigh syndrome; cancer; cardiomyopathy;
 KW ischemic disorder; inflammation; diabetes mellitus; fibrosis; mitosis;
 KW mesangial disorder; growth regulation; wound healing; T cell activation;
 KW immunosuppressant; primer; PCR; amplification; ss.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO9953036-A2.
 XX
 XX
 PD 21-OCT-1999.
 XX
 XX
 PF 13-APR-1999; 99WO-US08150.
 XX
 XX
 PR 14-APR-1998; 98US-0081784.
 XX
 XX (SUGEN-) SUGEN INC.
 PA
 PI Plowman G, Martinez R, Whyte D;
 XX
 XX WPI; 1999-611301/52.
 DR
 XX Novel kinase-related polypeptides used for the diagnosis and treatment
 PT of kinase-related diseases and disorders -
 XX
 PS Disclosure; Page 316; 387pp; English.
 XX

CC This sequence represents a degenerate oligonucleotide used to isolate the
 CC coding sequence for a novel STE20-related protein kinase. The invention
 CC relates to nucleic acid molecule encoding a kinase polypeptide selected
 CC from the kinases STK2, STK3, STK4, STK5, STK6, STK7, ZC1, ZC2, ZC3,
 CC ZC4, KHS2, SULU1, SULU3, GEK2, PAK4 and PAK5. The proteins are used to
 CC identify agonists and antagonists, and to raise antibodies. The
 CC polynucleotides are useful in gene therapy protocols. The polynucleotides,
 CC polypeptides, antibodies, antagonists and agonists may be used to treat
 CC diseases such as immune-related disorders and diseases (e.g. rheumatoid
 CC arthritis, artherosclerosis, chronic inflammatory bowel disease (e.g.
 CC Crohn's disease), asthma, osteoarthritis, psoriasis, atherosclerosis,
 CC rhinitis, autoimmunity, and organ transplantation, chronic inflammatory
 CC pelvic disease, multiple sclerosis, organ transplantation, myocardial
 CC infarction, cardiovascular disease, stroke, renal failure, oxidative
 CC stress-related neurodegenerative disorders (e.g. amyotrophic lateral
 CC sclerosis, Parkinson's disease and Leigh syndrome), cancer,
 CC cardiomyopathies, ischemic disorders, inflammatory disorders, diabetes
 CC mellitus, fibrotic and mesangial disorders. The proteins may also be
 CC useful for cell growth regulation (e.g. in wound healing), T cell
 CC activation, mitosis control, and as immunosuppressants.
 XX
 XX Sequence 23 BP; 3 A; 1 C; 8 G; 4 T; 7 other;
 SQ
 Query Match 4.6%; Score 16; DB 20; Length 23;
 Best Local Similarity 100.0%; Pred. No. 0.0023;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 322 GGGGARGTNTAYGARG 337
 Db 7 GGGGARGTNTAYGARG 22
 RESULT 3
 AAC84168
 ID AAC84168 standard; DNA; 885 BP.
 XX
 AC AAC84168;
 XX
 DT 19-MAR-2001 (first entry)
 XX
 DE Human OX2RH1 degenerate nucleotide sequence.
 XX
 KW OX2R protein; OX2RH1; OX2RH2; OX2RH3; OX2RH4; OX2RH1.2; ischaemia;
 KW antiinflammatory; cytoskeletal; neuroprotective; nootropic; human;
 KW antiarteriosclerotic; vasotropic; immunosuppressive; antirheumatic;
 KW antiarthritic; gene therapy; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200070045-A1.
 XX
 XX 23-NOV-2000.
 PD
 XX
 XX 11-MAY-2000; 2000WO-US12998.
 PF
 XX 13-MAY-1999; 99GB-0011123.
 PR
 XX 03-NOV-1999; 99GB-0025989.
 XX
 XX (MEDI-) MEDICAL RES COUNCIL.
 PA (SCHE) SCHERING CORP.
 XX
 XX Barclay AN, Brown MH, Gorman DM, Lanier LL, Wright GJ;
 PI Cherwinski H, Phillips JH, Hoek RM, Sedgwick JD;
 XX
 XX WPI; 2001-016233/02.
 DR P-PSDB; AAB48011.
 DR
 XX Mammalian OX2R proteins and DNA sequences useful for modulating the
 PT physiology and development of a cell -
 PT
 XX Disclosure; Page 35; 142pp; English.
 PS
 XX The invention relates to rodent or primate OX2R proteins, especially
 CC

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2002, 02:17:05 ; Search time 301 seconds
(without alignments)
2626.084 Million cell updates/sec

Title: US09897438BK-2

Perfect score: 351

Sequence: 1 GARCARTGYGCNACNATNAT.....AYGARGCNTGYTGCCNVTN 351

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 112599159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_101002.*

1: /SIDS2/cgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/cgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/cgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/cgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/cgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/cgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/cgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/cgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/cgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/cgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/cgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/cgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/cgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/cgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/cgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/cgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/cgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/cgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/cgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/cgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/cgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/cgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/cgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/cgdata/geneseq/geneseq-emb1/NA2002.DAT.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16	4.6	23	AAV42004	Receptor tyrosine
2	16	4.6	23	AAZ40496	Degenerate primer
3	16	4.6	885	AAC84168	Human OX2RH1 degen
4	16	4.6	1044	AAC84178	Human OX2RH1.2 deg
5	16	4.6	2379	AAH74766	Degenerate nucleot
6	15	4.3	426	AAAC91387	Human zaipha31 deg
7	15	4.3	696	AAAS13924	Degenerate DNA #1
8	15	4.3	696	ABK15334	Human helical prot
9	15	4.3	729	AAAS13926	Degenerate DNA #2

10	15	4.3	729	24	ABK15336	Human helical prot
11	15	4.3	987	22	AAF28673	Degenerate human z
12	15	4.3	1155	20	AAZ32748	Human protease-act
13	15	4.3	2019	22	AAH28471	Degenerate nucleot
14	15	4.3	2277	19	AAV13834	Homo sapiens ambig
15	15	4.3	2277	19	AAV05370	Human telomerase p
16	14	4.0	20	17	AAT45730	Primer PR2 used to
17	14	4.0	705	20	AAH08719	z219a degenerate p
18	14	4.0	1026	24	AAAS18143	Mouse DCRS10 rever
19	14	4.0	1134	22	AAAD09356	Human beta-1,3-gal
20	14	4.0	1218	20	AAH85600	Degenerate DNA enc
21	14	4.0	1584	22	AAF90613	Secretin-like rece
22	14	4.0	1629	22	AAF90614	Secretin-like rece
23	14	4.0	1695	24	AAAS18141	Human DCRS10 rever
24	14	4.0	1728	22	AAH28626	Human Kunitz domai
25	14	4.0	1728	24	ABK14096	Human zku8 degene
26	14	4.0	2094	24	AAAD24223	Murine zcytor14 cy
27	14	4.0	2094	24	AAAS18133	Mouse DCRS7 revers
28	13	3.7	17	18	AAT85341	Spider silk protei
29	13	3.7	20	14	AAQ40096	Thrombin-inhibiti
30	13	3.7	20	16	AAQ4020	Glutathione-indepe
31	13	3.7	23	15	AAQ70814	PCR primer used fo
32	13	3.7	23	16	AAQ92949	Saccharomyces cere
33	13	3.7	51	24	AAAD38244	Human zsig33-gamma
34	13	3.7	51	24	AAD25765	Human zsig33-gamma
35	13	3.7	351	22	AAAF83679	Human zsig33 degen
36	13	3.7	390	22	AAAD08417	Human DNAX cytokin
37	13	3.7	450	19	AAV45445	Human chemokine zS
38	13	3.7	669	20	AAAG1299	Human 2-19 protein
39	13	3.7	693	22	AAAD06415	Human DNAX cytokin
40	13	3.7	693	22	AAAF83736	Human cytokine rec
41	13	3.7	696	24	AAAD2922	Human soluble IL-2
42	13	3.7	729	24	AAAS18139	Mouse DCRS9 revers
43	13	3.7	747	22	AAAD08411	Human DNAX cytokin
44	13	3.7	783	21	AAZ52250	Human stomach prot
45	13	3.7	804	24	ABK48294	Degenerate DNA enc

ALIGNMENTS

RESULT 1

AAV42004
ID AAV42004 standard; DNA; 23 BP.

XX AAV42004;

AC AAV42004;

XX 12-OCT-1998 (first entry)

DT Receptor tyrosine kinase ROS-based primer ROS1.

XX Receptor tyrosine kinase; LMR1; LMR2; LMR3; signal transduction;

KW cancer; cancer; neurodegenerative disorder; gene therapy; PCR;

KW primer; ROS; ss.

XX Synthetic.

OS Synthetic.

XX WO9822507-A2.

PN 28-MAY-1998.

XX 21-NOV-1997; 97WO-US22526.

XX 22-NOV-1996; 96US-0031675.

PR (SUGEN-) SUGEN INC.

XX Joho KE, Plowman GD;

XX WPI; 1998-312419/27.

XX New isolated receptor tyrosine kinase genes - which are expressed in
PT neuronal tissues and tumour cells, useful as targets for

Query Match 4.6%; Score 16; DB 3; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.00046;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 322 GNGARGTNTAYGARG 337
|||||
Db 7 GNGARGTNTAYGARG 22

RESULT 2

US-09-053-866-3
; Sequence 3, Application US/09053866
; Patent No. 611075
; GENERAL INFORMATION:
; APPLICANT: Xu, Wenfeng
; APPLICANT: Presnell, Scott R.
; APPLICANT: Yee, David P.
; APPLICANT: Foster, Donald C.
; TITLE OF INVENTION: PROTEASE-ACTIVATED RECEPTOR
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053.866
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Leith, Debra K
; REGISTRATION NUMBER: 32,619
; REFERENCE/DOCKET NUMBER: 98-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6674
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1155 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-053-866-3

Query Match 4.3%; Score 15; DB 3; Length 1155;
Best Local Similarity 100.0%; Pred. No. 0.00082;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 145 TAYWSNGAYCCNWSN 159
|||||
Db 919 TAYWSNGAYCCNWSN 933

RESULT 3

US-09-479-130-3
; Sequence 3, Application US/09479130
; Patent No. 6436400
; GENERAL INFORMATION:
; APPLICANT: Xu, Wenfeng
; APPLICANT: Presnell, Scott R.
; APPLICANT: Yee, David P.

; APPLICANT: Foster, Donald C.
; TITLE OF INVENTION: PROTEASE-ACTIVATED RECEPTOR
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/479.130
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Leith, Debra K
; REGISTRATION NUMBER: 32,619
; REFERENCE/DOCKET NUMBER: 98-10
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6674
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1155 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-479-130-3

Query Match 4.3%; Score 15; DB 4; Length 1155;
Best Local Similarity 100.0%; Pred. No. 0.00082;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 145 TAYWSNGAYCCNWSN 159
|||||
Db 919 TAYWSNGAYCCNWSN 933

RESULT 4

US-08-676-967-2
; Sequence 2, Application US/08676967
; Patent No. 5747317
; GENERAL INFORMATION:
; APPLICANT: COLLINS, KATHLEEN
; TITLE OF INVENTION: Human Telomerase
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Science & Technology Law Group
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/676.967
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:

NAME: Osman Ph.D., Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UCB96-055
TELEPHONE: (415)343-4341
TELEFAX: (415)343-4342
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2277 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-08-676-967-2

Query Match 4.3%; Score 15; DB 1; Length 2277;
Best Local Similarity 100.0%; Pred. No. 0.00066;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 294 YCARTGGAARCARGA 308
|||||
DB 2106 YCARTGGAARCARGA 2120

RESULT 5
US-08-676-974-2
Sequence 2, Application US/08676974
Patent No. 5770422
GENERAL INFORMATION:
APPLICANT: COLLINS, KATHLEEN
TITLE OF INVENTION: Human Telomerase
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Science & Technology Law Group
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/676,974
FILING DATE:

CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UCB96-055
TELEPHONE: (415)343-4341
TELEFAX: (415)343-4342
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2277 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-08-676-974-2

Query Match 4.3%; Score 15; DB 1; Length 2277;
Best Local Similarity 100.0%; Pred. No. 0.00066;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 294 YCARTGGAARCARGA 308
|||||
DB 2106 YCARTGGAARCARGA 2120

RESULT 6
US-09-098-487-2
Sequence 2, Application US/09098487
Patent No. 5917025
GENERAL INFORMATION:
APPLICANT: COLLINS, Kathleen
TITLE OF INVENTION: Human Telomerase
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Science & Technology Law Group
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/098,487
FILING DATE:

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UCB96-055
TELEPHONE: (415)343-4341
TELEFAX: (415)343-4342
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2277 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-09-098-487-2

Query Match 4.3%; Score 15; DB 2; Length 2277;
Best Local Similarity 100.0%; Pred. No. 0.00066;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 294 YCARTGGAARCARGA 308
|||||
DB 2106 YCARTGGAARCARGA 2120

RESULT 7
US-08-418-893D-14/c
Sequence 14, Application US/08418893D
Patent No. 5559220
GENERAL INFORMATION:
APPLICANT: ROESSLER, PAUL G
APPLICANT: OHLROGE, JOHN B
TITLE OF INVENTION: GENE THAT ENCODES ACETYL-COENZYME A
TITLE OF INVENTION: CARBOXYLASE FROM CYCLOTELLA CRYPTICA
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: NATIONAL RENEWABLE ENERGY LABORATORY
STREET: 1617 Cole Blvd.
CITY: Golden
STATE: CO
COUNTRY: USA
ZIP: 80401-3393
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/418,893D

```
; FILING DATE: April 7, 1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/104,938
; FILING DATE: September 14, 1993
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: O'CONNOR, EDNA
; REGISTRATION NUMBER: 29,252
; REFERENCE/DOCKET NUMBER: MRI/NREL IR# 92-48CON
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303-231-1000
; TELEFAX: 303-231-1098
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-418-893D-14

Query Match      4.0%; Score 14; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 CAYGNAAYGCNGT 35
Db 14 CAYGNAAYGCNGT 1

RESULT 8
US-09-167-513-8
; Sequence 8, Application US/09167513
; Patent No. 6388064
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Blumberg, Hal
; TITLE OF INVENTION: A HUMAN 2-19 PROTEIN HOMOLOGUE, Z219A
; FILE REFERENCE: 97-63
; CURRENT APPLICATION NUMBER: US/09/167,513
; CURRENT FILING DATE: 1998-10-06
; EARLIER APPLICATION NUMBER: US 60/061,712
; EARLIER FILING DATE: 1997-10-06
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 705
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1)...(705)
; OTHER INFORMATION: Z219a Degenerate polynucleotide sequence
; OTHER INFORMATION: N is any nucleotide
; US-09-167-513-8

Query Match      4.0%; Score 14; DB 4; Length 705;
Best Local Similarity 100.0%; Pred. No. 0.0064;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 AYCNAARAAYAA 186
Db 632 AYCNAARAAYAA 645

RESULT 9
US-08-859-183-3
; Sequence 3, Application US/08859183
; Patent No. 5827731
```

```
; GENERAL INFORMATION:
; APPLICANT: Friedrich, Thomas
; APPLICANT: Bialojan, Siegfried
; APPLICANT: Bollschweiler, Claus
; APPLICANT: Kuenast, Christoph
; TITLE OF INVENTION: No. 5827731el thrombin-inhibitory protein
; TITLE OF INVENTION: from ticks
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Keil & Weinkeuf
; STREET: 1101 Connecticut Avenue
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.2 Mb storage
; COMPUTER: IBM AT-compatible, 80486 processor
; OPERATING SYSTEM: MS-DOS version 6.0
; SOFTWARE: WordPerfect version 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/859,183
; FILING DATE: 20-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/415,999
; FILING DATE: 04-APR-1995
; APPLICATION NUMBER: PCT/EP92/02179
; FILING DATE: 21-SEP-1992
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-859-183-3

Query Match      3.7%; Score 13; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 185 AYACNGCNGAYTG 197
Db 5 AYACNGCNGAYTG 17

RESULT 10
US-08-447-500-13
; Sequence 13, Application US/08447500
; Patent No. 5627064
; GENERAL INFORMATION:
; APPLICANT: Hoeckstra, Merl F.
; TITLE OF INVENTION: PROTEIN KINASES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East, Suite 500
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,500
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/008,001
; FILING DATE:
```

```
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell Ph.D., John R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD-2458
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: Protein Kinase
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..23
; OTHER INFORMATION: /note= "Bases designated N at
; positions 3, 6, 9, 12 and 18 are Inosine."
;
US-08-447-500-13

Query Match 3.7%; Score 13; DB 1; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GGNWSNGGNSNT 133
Db 1 GGNWSNGGNSNT 13

RESULT 11
US-08-454-097-13
; Sequence 13, Application US/08454097
; Patent No. 5686412
; GENERAL INFORMATION:
; APPLICANT: Hoekstra, Merl F.
; TITLE OF INVENTION: Protein Kinases
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,097
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/185,359
; FILING DATE: 21-JAN-1994
; APPLICATION NUMBER: US 08/008,001
; FILING DATE: 21-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,783
; FILING DATE: 03-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5686412and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31853
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 13:
```

```
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: Protein Kinase
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..23
; OTHER INFORMATION: /note= "Bases designated N at
; positions 3, 6, 9, 12 and 18 are Inosine."
;
US-08-454-097-13

Query Match 3.7%; Score 13; DB 1; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GGNWSNGGNSNT 133
Db 1 GGNWSNGGNSNT 13

RESULT 12
US-08-453-866-13
; Sequence 13, Application US/08453866
; Patent No. 5756289
; GENERAL INFORMATION:
; APPLICANT: Hoekstra, Merl F.
; TITLE OF INVENTION: PROTEIN KINASES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Spensley Horn Jubas & Lubitz
; STREET: 1880 Century Park East, Suite 500
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,866
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/008,001
; FILING DATE: 20-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Wetherell Ph.D., John R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD-2458
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 455-5100
; TELEFAX: (619) 455-5110
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: Protein Kinase
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..23
; OTHER INFORMATION: /note= "Bases designated N at
; positions 3, 6, 9, 12 and 18 are Inosine."
;
US-08-453-866-13
```

Query Match 3.7%; Score 13; DB 1; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 GGNWSNGNWSNT 133
|||||
Db 1 GGNWSNGNWSNT 13

RESULT 13

US-08-185-359-13
; Sequence 13, Application US/08185359
; Patent No. 6060296
; GENERAL INFORMATION:
; APPLICANT: Hoekstra, Merl F.
; TITLE OF INVENTION: Protein Kinases
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,359
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/008,001
; FILING DATE: 21-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/728,783
; FILING DATE: 03-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6060296and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31853
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: Protein Kinase
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..23
; OTHER INFORMATION: /note= "Bases designated N at
; positions 3, 6, 9, 12 and 18 are Inosine."
US-08-185-359-13

Query Match 3.7%; Score 13; DB 3; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 GGNWSNGNWSNT 133
|||||
Db 1 GGNWSNGNWSNT 13

RESULT 14

US-08-457-192-2
; Sequence 2, Application US/08457192
; Patent No. 5811514
; GENERAL INFORMATION:
; APPLICANT: Bard, Frederique
; APPLICANT: vednock, Theodore A.
; APPLICANT: Keim, Pamela S.
; TITLE OF INVENTION: NOVEL INTEGRIN BLOCKING MOLECULE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Hourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,192
; FILING DATE: HEREWITH
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/104,210
; FILING DATE: 09-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 15270-001330
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 72 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-457-192-2

Query Match 3.7%; Score 13; DB 1; Length 72;
Best Local Similarity 100.0%; Pred. No. 0.089;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 148 WSNAYCCNWSNA 160
|||||
Db 37 WSNAYCCNWSNA 49

RESULT 15

US-09-124-541-10
; Sequence 10, Application US/09124541A
; Patent No. 6229666
; GENERAL INFORMATION:
; APPLICANT: Morris Ph.D., Roy O.
; TITLE OF INVENTION: A CYTOKININ OXIDASE
; FILE REFERENCE: UMO1490
; CURRENT APPLICATION NUMBER: US/09/124,541A
; CURRENT FILING DATE: 1998-07-29
; EARLIER APPLICATION NUMBER: 60/054,268
; EARLIER FILING DATE: 1997-07-30
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 1602
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:

```
NAME/KEY: variation
LOCATION: (6)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (9)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (12)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (21)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (24)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (27)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (30)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (33)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (36)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (42)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (48)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (54)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (57)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (60)
OTHER INFORMATION: a,c,g or t
FEATURE:
NAME/KEY: variation
LOCATION: (63)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (66)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (69)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (72)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
```

```
LOCATION: (75)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (78)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (81)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (90)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (93)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (96)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (99)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (105)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (108)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (111)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (114)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (117)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (120)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (123)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (126)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (129)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (135)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (141)
OTHER INFORMATION: a,g,c or t
FEATURE:
NAME/KEY: variation
LOCATION: (144)
```

```
; OTHER INFORMATION: a,g,c or t
; FEATURE:
; NAME/KEY: variation
; LOCATION: (147)
; OTHER INFORMATION: a,g,c or t
; FEATURE:
; NAME/KEY: variation
; LOCATION: (153)
; OTHER INFORMATION: a,g,c or t
; FEATURE:
; NAME/KEY: variation
; LOCATION: (159)
; OTHER INFORMATION: a,c,g or t
; FEATURE:
; NAME/KEY: variation
; LOCATION: (162)
; OTHER INFORMATION: a,g,c or t
; FEATURE:
; NAME/KEY: variation
; LOCATION: (165)
; OTHER INFORMATION: a,g,c or t
; FEATURE:
; NAME/KEY: variation
; LOCATION: (168)
; OTHER INFORMATION: a,g,c or t
; FEATURE:
; NAME/KEY: variation
; LOCATION: (171)
; OTHER INFORMATION: a,g,c or t
; FEATURE:
; NAME/KEY: variation
; LOCATION: (174)
; OTHER INFORMATION: a,g,c or t
; FEATURE:
; NAME/KEY: variation
; LOCATION: (177)
; OTHER INFORMATION: a,g,c or t
```

Query Match 3.7%; Score 13; DB 4; Length 1602;

Best Local Similarity 100.0%; Pred. No. 0.032;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 96 NGCNWSNGTNTN 108

Db 1050 NGCNWSNGTNTN 1062

Search completed: November 7, 2002, 04:59:56
Job time : 72 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 6, 2002, 23:36:55 ; Search time 2767 Seconds
(without alignments)
3691.755 Million cell updates/sec

Title: US09897438BK-2
Perfect score: 351
Sequence: 1 GARGCARTGYGNCACATNAT.....AYGARGCNCYTGCGGCNTN 351-

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 20000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB	ID	Description
1	249.4	71.1	11673	6	AX305377	AX305377 Sequence
2	249.4	71.1	11673	10	MMU24703	U24703 Mus musculus
3	240.6	68.5	11187	10	AB049473	AB049473 Rattus no
4	233.8	66.6	11580	6	AX410790	AX410790 Sequence
5	233.8	66.6	11580	9	HSU79716	U79716 Human reeli
6	233.8	66.6	11580	11	G30936	G30936 SWS2926 Er
7	233.8	66.6	11580	11	G30938	G30938 SWS3176 Er
8	172.8	49.2	10634	5	AF090441	AF090441 Gallus ga
c 9	96.6	27.5	163985	10	AC121878	AC121878 Mus muscu
10	96.6	27.5	183156	2	AC023062	AC023062 Mus muscu
11	92.6	26.4	126130	2	AC095877	AC095877 Rattus no
c 12	92.6	26.4	202764	2	AC128022	AC128022 Rattus no
c 13	91.6	26.1	183641	2	AC124933	AC124933 Rattus no
c 14	90.2	25.7	93163	9	HSAC000121	AC000121 Human BAC
c 15	53.4	15.2	185996	2	AC041023	AC041023 Homo sapi
16	49.6	14.1	158	4	AF232904	AF232904 Bos tauru
17	48	13.7	125020	9	AF429315	AF429315 Homo sapi
c 18	47	13.4	125020	9	AF429315	AF429315 Homo sapi
19	45.2	12.9	146383	2	AC116367	AC116367 Oryza sat
20	43.8	12.5	161034	9	AC015771	AC015771 Homo sapi
21	43.8	12.5	161707	2	AC015772	AC015772 Homo sapi
c 22	43.6	12.4	115666	2	AC105744	AC105744 Oryza sat
c 23	42.6	12.1	115666	2	AC105744	AC105744 Oryza sat
c 24	42	12.0	68341	2	AC105584	AC105584 Rattus no
c 25	41.8	11.9	173926	2	MM454K24	AL160232 Mus muscu
26	41.6	11.9	104695	9	HSUJ96108	AL121879 Human DNA
c 27	41.6	11.9	162355	2	AC092957	AC092957 Homo sapi
c 28	41.6	11.9	176186	9	AC006925	AC006925 Homo sapi
c 29	41.6	11.9	188688	2	AL833801	AL833801 Mus muscu
c 30	41.2	11.7	153645	9	AC025426	AC025426 Homo sapi
31	41.2	11.7	165110	9	AL450311	AL450311 Human DNA
c 32	40.6	11.6	33556	2	AC020474	AC020474 Drosophil
33	40.6	11.6	172105	3	AC104625	AC104625 Drosophil
c 34	40.6	11.6	183334	2	AC106174	AC106174 Rattus no
c 35	40.6	11.6	302915	3	AE003489	AE003489 Drosophil
c 36	40.2	11.5	1141	6	AX083744	AX083744 Sequence
37	40.2	11.5	1246	6	AX164174	AX164174 Sequence
c 38	39.8	11.3	146383	2	AC116367	AC116367 Oryza sat
c 39	39.6	11.3	173341	2	AC021954	AC021954 Homo sapi
c 40	39.4	11.2	157197	2	AC099042	AC099042 Oryza sat
c 41	39.4	11.2	192195	2	AC099620	AC099620 Mus muscu
c 42	39.2	11.2	3461	3	DROM1RADH	M60998 Drosophila
c 43	39.2	11.2	7218	6	I66494	I66494 Sequence 14
c 44	39.2	11.2	156206	2	AC105500	AC105500 Rattus no
c 45	39.2	11.2	293184	2	AC079487	AC079487 Mus muscu

ALIGNMENTS

RESULT 1
AX305377
LOCUS AX305377 11673 bp DNA linear PAT 11-DEC-2001
DEFINITION Sequence 128 from Patent WO0188188.
ACCESSION AX305377
VERSION AX305377.1 GI:17644926
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scturognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Ishikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y.
TITLE Method for examining ischemic conditions
JOURNAL Patent: WO 0188188-A 128 22-NOV-2001;

Pred. No. is the number of results predicted by chance to have a


```
Query Match 71.1%; Score 249.4; DB 10; Length 11673;
Best Local Similarity 57.7%; Pred. No. 2.6e-56;
Matches 202; Conservative 79; Mismatches 69; Indels 0; Gaps 0;

QY 1 GARCARTGYGNCNATNATGTCAYGGNAAYGCNCTNACNTTYTGYGARCNCNTAYGGNCN 60
  ||:||||:| || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 970 GAGCAGTGTGCACCATCATGTCATGGCAATGCTCTACCTCTCTGTGAGCCCTACGGCCCT 1029
  ||:||||:| || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 61 MNGARTYNACNACNATNGTYTNAAYACNACNACNCGNWSNGNTYNTNCARTTYWSNATN 120
  ||:||||:| || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1030 CGAGAGTGCACCATCATGCTGACACACACACAGCATGCTCTCCAGCTTTCATT 1089
  ||:||||:| || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 121 GGNWSNGNWSNTGYMNTTYWSNTYWSNGAYCCNWSNATNACNCTNWSNTAYGCNNAAR 180
  ||:||||:| || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1090 GGGTCAGGATCATGTCGATTAGTTACTCTGACCCAGCATCACTGTGTGCATACGCCAAG 1149
  ||:||||:| || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 181 AAYAAACNCGNAYTGATNARYTNGARARATNNGNCGNWSNAYAGTWNWSNACN 240
  ||:||||:| || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1150 AACAAACCGTGAATGGATTCAGTTCAGTGAGAAAATTAGAGCCCTTCCAAATGTGAGACA 1209
  ||:||||:| || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 241 GTNATNCAYATNTYNTAYTNCNARGARGCNAARGNGARWSNGTNCARTTYCARTGG 300
  ||:||||:| || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1210 GTATCCACATCTGTACCTCCCGAGAGAACCCAAAGGAGAGCGGTGCAGTTCAGTGG 1269
  ||:||||:| || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 301 AARCAGAYWSNTYNTMGNGTNGGARGTNTAYGARGCNTGYTGGGCTNYT 350
  ||:||||:| || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1270 AACAGCAGACGCTGCCAGTGGGTGAGGTGTATGAGGCGCTGCTGGGCGCT 1319
  ||:||||:| || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 3
AB049473
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

FEATURES
source
1. .11187
  /organism="Rattus norvegicus"
  /db_xref="taxon:10116"
  /tissue_type="cerebellum"
  /dev_stage="juvenile"
  1. .11187
  /gene="Reln"
  341. .10729
  /gene="Reln"
  /codon_start=1
  /product="reelin"
  /protein_id="BAB/8470.1"
  /db_xref="GI:17221618"
  /translation="MERGCAWPTLVLLVALLLALTLARAAATGYPRFSPFFFLCTH
HGELEGDQGEVLI SLHAGNPYVVPQGVHVTJSTFFDGLVLTGLYTSIQS
SOSIGSSAFGFCINSHDHOFMCSVAVSHVSLPTNLSFVWAPAGTCVNFN
ATATHRGQVIFDQALQCEQCAPTEAYSHLAEIHSDSVILRDFDQSYHOLELN
NWAECNSDTEGEQGTIMHGNVATEPEPYPRELTYLNTTATSVLQFSIGSGSCR
FYSDFPSTIIVSAKNNTADWIOLEKIRAPSNVSTIIHILYLPEDAKGENVQFQWQDS

Rattus norvegicus 11187 bp mRNA linear ROD 01-DEC-2001
Rattus norvegicus mRNA for reelin, complete cds.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1
Kikkawa,S. and Terashima,T.
rat reelin (Reln) complete CDS
Published Only in Database (2001)
2 (bases 1 to 11187)
Kikkawa,S. and Terashima,T.
Direct Submission
Submitted (29-SEP-2000) Satoshi Kikkawa, Kobe University School of
Medicine, Department of Anatomy; Chuo-ku, Kusunoki-cho 7-5-1, Kobe,
Hyogo 650-0017, Japan (E-mail:skikkawa@med.kobe-u.ac.jp,
Tel:81-78-382-5325, Fax:81-78-382-5328)
Location/Qualifiers
1. .11187
  /organism="Rattus norvegicus"
  /db_xref="taxon:10116"
  /tissue_type="cerebellum"
  /dev_stage="juvenile"
  1. .11187
  /gene="Reln"
  341. .10729
  /gene="Reln"
  /codon_start=1
  /product="reelin"
  /protein_id="BAB/8470.1"
  /db_xref="GI:17221618"
  /translation="MERGCAWPTLVLLVALLLALTLARAAATGYPRFSPFFFLCTH
HGELEGDQGEVLI SLHAGNPYVVPQGVHVTJSTFFDGLVLTGLYTSIQS
SOSIGSSAFGFCINSHDHOFMCSVAVSHVSLPTNLSFVWAPAGTCVNFN
ATATHRGQVIFDQALQCEQCAPTEAYSHLAEIHSDSVILRDFDQSYHOLELN
NWAECNSDTEGEQGTIMHGNVATEPEPYPRELTYLNTTATSVLQFSIGSGSCR
FYSDFPSTIIVSAKNNTADWIOLEKIRAPSNVSTIIHILYLPEDAKGENVQFQWQDS
```

QY	181	AAAYACNGCNGAYTGGATNCARTNGTNGARARATNMNGNCNCCNWSNAAYGTNWSNACN	240
Db	1211	AACAATAGTGGCGATTGGATTGAGTCGAGAGATTAGAGCCCTTCCAACGTAAGCACCC	1270
QY	241	GTNATNCAYATNTYATYTNCCNGARGCNAARGNGGARGWSNGTNCARTTYCARTGG	300
Db	1271	ATCATCCACATCTCTACTCTCTGAGGACGCCAAGGAGGAGAGTCCAGTTCAGTGG	1330
QY	301	AARCAAGAYWSYTNMNGTNGGNGARGTNTAYGARGCNTGYTGGCNYT	350
Db	1331	AAACAGGACAGCTGCATGTGGTGGGTGAGGTGACGAGGCTGCTGGGCCCT	1380
RESULT 4			
LOCUS	AX410790	11580 bp	DNA
DEFINITION	Sequence 3437 from Patent WO0229103.	linear	PAT 14-JUN-2002
ACCESSION	AX410790		
VERSION	AX410790.1	GI:21443495	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
JOURNAL	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
FEATURES	Location/Qualifiers		
source	1. .11580		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/note="EMBL/GenBank Accession No. U79716"		
BASE COUNT	3014 a 2696 c 2753 g 3116 t	1 others	
ORIGIN			
Query Match	66.6%;	Score 233.8;	DB 6;
Best Local Similarity	56.0%;	Pred. No. 4.5e-52;	
Matches	196;	Conservative 74;	Mismatches 80;
		Indels 0;	Gaps 0;
QY	1	GARCARTGGNACNATNATGAYGNGNAYGNGTNCARTTYTGARGCNCNTAYGGNCCN	60
Db	860	GAACAGTGGCGGATATGATCGATGCGCAATCCGTCACCTTCTGTGAACCATATGCCCA	919
QY	61	MNGARYTNACNACNATGYTNAAYACNACNACNGNWSNGTNTNCARTTYWSNATN	120
Db	920	CGAGACTGATTACACAGGCTTAATACACACAGCTTCTGCTCCCAATTTCCATT	979
QY	121	GGNWSNGNWSNTGYMGTNTYWSNTAYWSNGAYCCNWSNATNACNNGTNSNTAYGCNNAAR	180
Db	980	GGTCAAGTTTCATGTCGCTTTAGTTATTTCAGACCCAGCATCATCTGTTATATGCCAAG	1039
QY	181	AAAYACNGCNGAYTGGATNCARTNGTNGARARATNMNGNCNCCNWSNAAYGTNWSNACN	240
Db	1040	AATACTCTGGCGATTCAGTTCAGTAGAGAAAATTAGAGCCCTTCCAATGTCCAGACA	1099
QY	241	GTNATNCAYATNTYATYTNCCNGARGCNAARGNGGARGWSNGTNCARTTYCARTGG	300
Db	1100	ATCATCCATATCTCTACTCTCTGAGGACGCCAAGGAGGAGATGTCCAATTCAGTGG	1159
QY	301	AARCAAGAYWSYTNMNGTNGGNGARGTNTAYGARGCNTGYTGGCNYT	350
Db	1160	AAGCAGGAAAATCTCTGCTAGGTGAGTGAAGTGTATGAAGCCTGCTGGGCCCT	1209
RESULT 5			
LOCUS	HSU79716	11580 bp	mrna
DEFINITION	Human reelin (RELN) mrna, complete cds.	linear	PRI 25-FEB-1997
ACCESSION	U79716		
VERSION	U79716.1	GI:1743884	
KEYWORDS			

SOURCE	Homo sapiens.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE	1 (bases 1 to 11580)
JOURNAL	Desilva,U., D'Arcangelo,G., Braden,V.V., Chen,J., Miao,G.G.,
MEDLINE	Curran,T. and Green,E.D.
PUBMED	The human reelin gene: isolation, sequencing, and mapping on
REFERENCE	chromosome 7
AUTHORS	Genome Res. 7 (2), 157-164 (1997)
TITLE	97202106
JOURNAL	2 (bases 1 to 11580)
MEDLINE	Desilva,U., D'Arcangelo,G., Braden,V.V., Chen,J., Miao,G.G.,
PUBMED	Curran,T. and Green,E.D.
REFERENCE	Direct Submission
AUTHORS	Submitted (26-NOV-1996) National Center for Human Genome Research,
TITLE	National Institutes of Health, 49 Convent Drive, MSC4431, Bethesda,
JOURNAL	MD 20892, USA
FEATURES	Location/Qualifiers
source	1. .11580
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="7"
	1. .11580
	/gene="RELN"
	176..10558
	/gene="RELN"
	/codon_start=1
	/product="reelin"
	/protein_id="AAC51105.1"
	/db_xref="GI:1743885"
	/translation="MERSGAROTFLLALLGATLRARAAAGYVPRESPEFLCTHGG
	ELEGDQGEVLISLHAGNPTYYVQGEYHVHTISTSTFPDGLLVLTGYSTVSQASQ
	SIGGSAGFGIMSDHQFGNOFMCVSVASHVHPTTNLSFNTWAPAGTCVNFMT
	ATHRGQVIFKDALQCEQAGPTDVVPHLAIEHSDSIILRDDDFSYHQLQNPNI
	YVNCNCTGEQCGAIMEGNAVTECEPYGRELITTLGNTTASLVDSIGSGSCRS
	YSDPSIIVAKNSADWIOLEKIRAPSNVSTIHLVPEDAKGNVFOVKOENLIR
	VGEYECWALDNILIINSAHROVLEDSLDPDVTGNWLPFGATVHKSCQDGSNIY
	PHNGEGSEFNATRDVLDSTEDIQEWSSEFESQPTGWDVLGAVIGTEGCTIESGLS
	MVLKYSERKLCPTSMDTTGNLRFVFMGICDPNGSHENDIILAKIBGRHEIT
	LDLTLSYSGKVPVSLVYINPELOTPATKFCLOKNOHGNHNNVYDFVHLPVLS
	TSMHISQINLGCTHQPGNSVLESTNHSRSLNLTHTCELCPEICAGLPHSTVY
	SSNIYSGWNRITIPLPNAALTNRIRWROTGPLNGLNWAIDNVIQPSCLKSGRG
	QCTRHGCKDPGFGSPACEMASOTFPMFISESGSRLSSVHNFYSIRGAEVSGCV
	LASGKALVENKEGPROLITSEFLDSSOSREFLOFTLRGSKSVLSTRAPDQGEVLLH
	YSDNGITWLLKLEHYSLSLSTHEPRIISVELPDQAKFGQIFRWQPHSSOREDYWAI
	DEIIMTSVFNISLDFTNLVETVQSLGFLGNVQPCYGHWDITLCFTGDSKLASSMY
	VETQSMQIGASYMIQFSLVGGCGQKYPHMDNQVKLEYSTNHGLTWHLVQEECLSP
	SCOEFTSASIVHASEFTQWRVIVLLPQKTWSSATFRWSOSYTAQDEWALDSIYG
	QOCNMGSGHSCDHGICRCDOGYGTECHPEAALPSTINSDRENQNGWESDQEVIG
	GEIVKPEQCGVLSGSSLSYFSKAGKQLVSWDLDTSWDPVQFYIQLGSESASCNKP
	DSREBVLLQYSNNGGQIOWHLLAEIYFSDFSKPRVYLELPAAAKTCTCRFWQPVF
	SGEDYDQWAVDDIIILSEKOKIIPINPTLPQNFYKPAFDPMNQMSVWMLANEG
	MYNTECAATPSAMIKSGDGRFAVTRDITLTKPGVYLAQKLNIGCANQFSSTAPVL
	LOYSHDAGMSWFLVKEGYPASAGCGEGNSRELSEPTMYHTGDFEWTRTIIVPRS
	LASSKTRRWIOESSOKNVPFGLDGYISEPCCSYCSGHDICI SGVPCDLCYTA
	QCTCVSNVPHNEMDFEGKSLPLWKITGAQVGTGCGTINDKSLYFNGPREAR
	TVPLDTRNRLVQFYIQLGSGTSGITCICPKRNEGLIVQSDNGLHLLRLDEFP
	SFLEPQIISIDLPODAKTPATFRWQPHGKHSQAQWALDVLGIMDSSDGTQDFK
	DGSDIQANNYRTGGQVDICLSMDTALITENIGKPYATWDFWISASTFLQEMK
	SMGSKPFSNHSVQLQYSLNNGKDWLVTECPVPTIGCLHYTESSTIYTSERFQNM
	RTVYPLSTISPRTRFRVQANTVGDASWAIQNVLASGCPMWSGRCICDAGRCV
	CDRGFGGCVPVVPLPSILKDDPENGILHPLDLPVYGAERNGELTKSGTSLIKP
	BEGRLMLISDLCTNTMYVQFSRLFTAKSTPERSHSLIQLFSGSGITWHLMDFFY
	PQTNILFINVPLPYATQATNATRLWQPNYNNKKEIWIIVDDFIIDGNVNNVPMLL
	DTDFGPRDNWFFYPGGNIQLGYCPYSKGAPEDSAMVFSYNEVGHSTIRDLNVN
	ENTLIQEIINVGCTSSDSDAPVLEFSRDFGATWHLLLPLCYHSSHSVSLCSTEHH
	PSSTYAGTWMQNRVVRHKGKLCGSRFRVYQGFYAGSOPVWAIQNVYIGPOC
	REMNGOGSCINCTKICDPYSGPTCKISTKNPDLKDDPFGOLESDFLMSSGPK
	SRKCGILSSGNNLFFNEDGLMLMTOLDLSHARVQFFRHLGCGKGVDPDRSPVLL
	QYSLNGGLSWSLQELFNSNSVGRYIALEIPLKARSSTRLRWQPSSENGHFYSPW
	VIDQILIGGNSISGNTVLEDDFTTILDSRKWLLHPGGRKMPVCGSTGDAVLFTEKASTY

VVSTDAVAVNEDSLQIDFAASCSVTDSYAEILEYSVDLGLSWHPLVRDCLPTNVES
RYHOLIRLYSDFTNKCVRITLPLPPTRSOATFRWHQAPFDKQWADNVIYIGDG
CIDMCSHGRCIOGNCVCEOMGLYCDOPETSLTQKDNFNEPSSQWNLTVNGGK
LSTVCASVAGMALHFGSGCSRLITVVDLNLNAEPIQFYFMGCLITPNNRQGVLL
EYSVNGITWNLMELFYDIQSKAPGVNILLPPDAKEIATFRFWMQPRHDGLDNDWA
IDNVLSSGADQETVMDLTFSSAPQHERSPADAGPVGRIADFMEDKTSVNEHWL
FHDCTVERFCDSGVMLVPOCLPADPKCSGVSPVSFFPTKWKRTIYPLUESIA
QNDIHWQYSTDFGVSNLYVPOCLPADPKCSGVSPVSFFPTKWKRTIYPLUESIA
GNVPRREFOKYSMDOMADNDFYLGPCLDNCRGDCREOCICDPGSGPNCYLTH
TLKFLKERPDESEIKPDLWMSLEGGSTCTCEGLAEDIALYFGGTVRQAVQDLDL
RGAKFYQIRGSENNMTSCHRPICKRKEGLVDYSTDGTTWLLHEMDYQIYSVR
HDXLPELDEALNTLRWQPFVINGVSVGERAOWALDNLIGAEINPSQLVD
TFYDETTSEENMSFYNAVRTAGFCGNPSFLHWPNNKKDKTHNALSSRELIIQCY
MMQFKLVGCEATSCDGLHSMLEYTKDARSQWOLVQTOCLPSSNSIGCSPEOFHE
ATTYNSVSSSKRITITQLPDHVSSTATOPRWIKQGEETEKOSWADHWYIGACPKL
CSGGYCTTGACICDESEFGDCSVESHDLPYIKDNFESARTANWETIQGGVIG
SCGGCLAPTAHGSDFNGCQIHOAAATKPLDPLFRASKIMFYLOIGMSQTDSCNSDL
GPHAVDQVLLQYVSVNNGTHTWVIAHQPKDFQAOVSYNVPLEARMKGVLLRWMP
RHNGTGDQWALDHVEVLVSTYTRKQYMMNFSRQHGRLRHYNNRRSLRRYP*
BASE COUNT 3014 a 2696 c 2753 g 3116 t 1 others
ORIGIN

Query Match 66.6%; Score 233.8; DB 9; Length 11580;
Best Local Similarity 56.0%; Pred. No. 4.5e-52;
Matches 196; Conservative 74; Mismatches 80; Indels 0; Gaps 0;
QY 1 GARCARTGYGNACNATNATCCAYGNAAYCGNCTNACNTTYTGCGARCCNTAYGGNCCN 60
DB 860 GAACAGTGTGGCGGATATGTCATGCAATGCCGTACCCTCTGTGAACCATATGGCCCA 919
QY 61 MNGARYTNACNACNACNTGYTNAAYACNACNACNNGTNTYTCARTTYWSNATN 120
DB 920 CGAAGTGTATACACAGGCTTAAACACACAGCTTCTGTCTCCATTTCCATT 979
QY 121 GGNWSNGNSNTGYMNTYWSNTAYWSNGAYCCNWSNATNACNCTNWSNTAYGCNAAR 180
DB 980 GGGTCAGGTTTCCTGCTTGTATTCAGACCCAGCATCATCTGTATTATGCGCAAG 1039
QY 181 AAYAAACNCGNAYTCGATNCAATYNGARAATNNGNCCNWSNAYAGTWNACN 240
DB 1040 AATAACTCTCGGACTGGATTCAGTAGAGAAATAGACCCCTTCCATGTGAGCACA 1099
QY 241 GTNATNCAATNTYNTAYTNCNGARGCNAARGNGARWSNGTNCARTTYCARTGG 300
DB 1100 ATCATCCATATCTCTACCTTCTGAGGACCCCAAGGAGAGATGTCCATTTTCAGTGG 1159
QY 301 AARCAAGYNSNTNMNGNTNGNGARGTNTAYGARGCNTGYTGGGCGNYT 350
DB 1160 AAGCAGGAAATCTTCTGTAGTGAAGTGTATGAGGCTGCTGGGCGTT 1209

RESULT 6
G30936
LOCUS
DEFINITION G30936 11580 bp DNA linear STS 28-SEP-1998
SWSS2926 Eric D. Green Homo sapiens STS genomic, sequence tagged
site.
ACCESSION G30936
VERSION G30936.1 GI:1923209
KEYWORDS STS.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 11580)
Bouffard,G.G., Iyer,L.M., Idol,J.R., Braden,V.V., Cunningham,A.F.,
Weintraub,L.A., Mohr-Tidwell,R.M., Peluso,D.C., Fulton,R.S.,
Leckie,M.P. and Green,E.D.
A collection of 1814 human chromosome 7-specific STSs
Genome Res. 7 (1), 59-64 (1997)
TITLE
JOURNAL 97189344
MEDLINE 9037602
PUBMED
REFERENCE 2 (bases 1 to 11580)
AUTHORS Green,E.D.

TITLE
JOURNAL
COMMENT

Human chromosome 7 STSs (1997)
Unpublished (1997)
On Apr 3, 1997 this sequence version replaced gi:1706935.
Synonyms: RELN
GDB_DSEG: RELN
Contact: Eric D. Green
Genome Technology Branch
National Human Genome Research Institute/NIH
49 Convent Dr., MSC4431, Bldg. 49, Rm. 2A08, Bethesda, MD 20892
Tel: 3014020201
Fax: 3014024735
Email: egreen@hgrl.nih.gov
Primer A: CTCCTCTCTCTGAGGACGCCAA
Primer B: AAGCCAGTTGCCCTGTCTCCACTG
STS size: 183
PCR profile:
Presoak: 0 degrees C for 0.00 minute(s)
Denaturation: 92 degrees C for 0.17 minute(s)
Annealing: 55 degrees C for 1.00 minute(s)
Polymerization: 72 degrees C for 1.00 minute(s)
PCR Cycles: 35
Thermal Cycler: PerkinElmer 9600
Protocol:
Template: 30-100 ng
Primer: each 1 uM
dNTPs: each 200 uM
Taq Polymerase: 0.05 units/ul
Total Vol: 10 ul
Buffer:
MgCl2: 1.5 mM
KCl: 100 mM
Tris-HCl: 10 mM
NH4Cl: 5 mM
pH: 8.6

This STS was developed from sequence determined by another investigator. See GenBank record: U79716 For additional information about the NHGRI chromosome 7 mapping project, see http://www.nhgri.nih.gov/DIR/GRB/CHR7. Also see Genomics 11:548-64 (1991) [MUID=92128937].
FEATURES
source 1..11580
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="7"
gene 1..11580
/gene="RELN"
STS 1112..1294
/gene="RELN"
primer_bind 1112..1134
/gene="RELN"
primer_bind complement(1272..1294)
BASE COUNT 3014 a 2696 c 2753 g 3116 t 1 others
ORIGIN

Query Match 66.6%; Score 233.8; DB 11; Length 11580;
Best Local Similarity 56.0%; Pred. No. 4.5e-52;
Matches 196; Conservative 74; Mismatches 80; Indels 0; Gaps 0;
QY 1 GARCARTGYGNACNATNATCCAYGNAAYCGNCTNACNTTYTGCGARCCNTAYGGNCCN 60
DB 860 GAACAGTGTGGCGGATATGTCATGCAATGCCGTACCCTCTGTGAACCATATGGCCCA 919
QY 61 MNGARYTNACNACNACNTGYTNAAYACNACNACNNGTNTYTCARTTYWSNATN 120
DB 920 CGAAGTGTATACACAGGCTTAAACACACAGCTTCTGTCTCCATTTCCATT 979
QY 121 GGNWSNGNSNTGYMNTYWSNTAYWSNGAYCCNWSNATNACNCTNWSNTAYGCNAAR 180
DB 980 GGGTCAGGTTTCCTGCTTGTATTCAGACCCAGCATCATCTGTATTATGCGCAAG 1039

DB 15322 GGGGAGAGCGTGCAAGTTCAGTGGAAACAGGACAGCCGCGAGTGCGGTGAGGTGTATGAG 15263

[illegible]

Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokenko, S., Oguh, M., Okwuonu, G.,
 Oraqunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
 Rives, M., Rojas, A., Rojokan, I., Rolfe, M., Ruiz, S., Savery, G.,
 Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I.,
 Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstein, G., and Gibbs, R.
 Direct Submission
 Unpublished
 2 (bases 1 to 126130)
 Worley, K.C.
 Direct Submission
 Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 126130)
 Worley, K.C.
 Direct Submission
 Submitted (11-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 10, 2002 this sequence version replaced gi:20975941.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GDRS
 Center clone name: CH230-10G13
 ----- Summary Statistics
 Sequencing vector: Plasmid
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap: version 0.990329
 Consensus quality: 70934 bases at least Q40
 Consensus quality: 75985 bases at least Q30
 Consensus quality: 80731 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 59 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

12811 14180: contig of 1370 bp in length
 14181 14280: gap of unknown length
 16048 16048: contig of 1768 bp in length
 16148 16148: gap of unknown length
 17569 17569: contig of 1421 bp in length
 17570 17570: gap of unknown length
 17670 17670: contig of 1151 bp in length
 18920 18920: gap of unknown length
 20209 20209: contig of 1289 bp in length
 20309 20309: gap of unknown length
 21633 21633: contig of 1324 bp in length
 21733 21733: gap of unknown length
 23269 23269: contig of 1536 bp in length
 23369 23369: gap of unknown length
 24647 24647: contig of 1278 bp in length
 24747 24747: gap of unknown length
 25800 25800: contig of 1053 bp in length
 25900 25900: gap of unknown length
 27017 27017: contig of 1117 bp in length
 27117 27117: gap of unknown length
 29183 29183: contig of 2066 bp in length
 29283 29283: gap of unknown length
 30315 30315: contig of 1032 bp in length
 30415 30415: gap of unknown length
 32086 32086: contig of 1671 bp in length
 32186 32186: gap of unknown length
 33395 33395: contig of 1209 bp in length
 33495 33495: gap of unknown length
 35663 35663: contig of 2168 bp in length
 35763 35763: gap of unknown length
 38044 38044: contig of 2281 bp in length
 38144 38144: gap of unknown length
 39845 39845: contig of 1701 bp in length
 39945 39945: gap of unknown length
 41982 41982: contig of 2037 bp in length
 42082 42082: gap of unknown length
 43743 43743: contig of 1661 bp in length
 43843 43843: gap of unknown length
 45904 45904: contig of 2061 bp in length
 46004 46004: gap of unknown length
 47667 47667: contig of 1663 bp in length
 47767 47767: gap of unknown length
 49833 49833: contig of 2066 bp in length
 49933 49933: gap of unknown length
 51434 51434: contig of 1501 bp in length
 51534 51534: gap of unknown length
 53778 53778: contig of 2244 bp in length
 53878 53878: gap of unknown length
 55552 55552: contig of 1674 bp in length
 55652 55652: gap of unknown length
 57670 57670: contig of 2018 bp in length
 57770 57770: gap of unknown length
 59630 59630: contig of 1860 bp in length
 59730 59730: gap of unknown length
 61147 61147: contig of 1417 bp in length
 61247 61247: gap of unknown length
 63995 63995: contig of 2748 bp in length
 64095 64095: gap of unknown length
 66294 66294: contig of 2199 bp in length
 66394 66394: gap of unknown length
 68794 68794: contig of 2400 bp in length
 68894 68894: gap of unknown length
 71533 71533: contig of 2639 bp in length
 71633 71633: gap of unknown length
 73569 73569: contig of 1836 bp in length
 73694 73694: gap of unknown length
 77054 77054: contig of 3485 bp in length
 77154 77154: gap of unknown length
 79005 79005: contig of 1851 bp in length
 79105 79105: gap of unknown length
 81294 81294: contig of 2189 bp in length
 81394 81394: gap of unknown length
 83864 83864: contig of 2470 bp in length

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT


```

Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G., and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 183641)
Worley,K.C.
Direct Submission
Submitted (20-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 183641)
Worley,K.C.
Direct Submission
Submitted (31-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 24, 2002 this sequence version replaced gi:21490053.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KAVI
Center clone name: CH230-466N17
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 111724 bases at least Q40
Consensus quality: 120970 bases at least Q30
Consensus quality: 128458 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 88 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1107: contig of 1107 bp in length
* 1108: gap of unknown length
* 1208: contig of 1253 bp in length
* 2461: 2560: gap of unknown length
* 2561: 4004: contig of 1444 bp in length
* 4005: 4104: gap of unknown length
* 4105: 5382: contig of 1278 bp in length
* 5383: 5482: gap of unknown length
* 5483: 6874: contig of 1392 bp in length
* 6875: 6975: gap of unknown length
* 6975: 8071: contig of 1097 bp in length
* 8072: 8171: gap of unknown length
* 8172: 9478: contig of 1307 bp in length
* 9479: 9578: gap of unknown length
* 9579: 11061: contig of 1483 bp in length
* 11062: 11161: gap of unknown length
*
11162: 12742: contig of 1581 bp in length
12743: 12842: gap of unknown length
12843: 14504: contig of 1662 bp in length
14504: 14604: gap of unknown length
14604: 15734: contig of 1130 bp in length
15734: 15735: gap of unknown length
15735: 16958: contig of 1124 bp in length
16958: 17059: gap of unknown length
17059: 18704: contig of 1646 bp in length
18704: 18804: gap of unknown length
18804: 20420: contig of 1616 bp in length
20420: 20520: gap of unknown length
20520: 21648: contig of 1128 bp in length
21648: 21748: gap of unknown length
21748: 23177: contig of 1429 bp in length
23177: 23277: gap of unknown length
23277: 24760: contig of 1483 bp in length
24760: 24860: gap of unknown length
24860: 26759: contig of 1899 bp in length
26759: 26859: gap of unknown length
26859: 28299: contig of 1440 bp in length
28299: 28399: gap of unknown length
28399: 29592: contig of 1193 bp in length
29592: 29692: gap of unknown length
29692: 30870: contig of 1178 bp in length
30870: 30970: gap of unknown length
30970: 32642: contig of 1672 bp in length
32642: 32742: gap of unknown length
32742: 34386: contig of 1644 bp in length
34386: 34486: gap of unknown length
34486: 36064: contig of 1578 bp in length
36064: 36164: gap of unknown length
36164: 37398: contig of 1234 bp in length
37398: 37498: gap of unknown length
37498: 39608: contig of 2110 bp in length
39608: 39708: gap of unknown length
39708: 41480: contig of 1772 bp in length
41480: 41580: gap of unknown length
41580: 42759: contig of 1179 bp in length
42759: 42859: gap of unknown length
42859: 44711: contig of 1852 bp in length
44711: 44811: gap of unknown length
44811: 46208: contig of 1397 bp in length
46208: 46308: gap of unknown length
46308: 47667: contig of 1459 bp in length
47667: 47867: gap of unknown length
47867: 49227: contig of 1360 bp in length
49227: 49327: gap of unknown length
49327: 50797: contig of 1470 bp in length
50797: 50897: gap of unknown length
50897: 51928: contig of 1031 bp in length
51928: 52028: gap of unknown length
52028: 53105: contig of 1077 bp in length
53105: 53205: gap of unknown length
53205: 54546: contig of 1341 bp in length
54546: 54646: gap of unknown length
54646: 55820: contig of 1174 bp in length
55820: 55920: gap of unknown length
55920: 57635: contig of 1715 bp in length
57635: 57735: gap of unknown length
57735: 58942: contig of 1207 bp in length
58942: 59042: gap of unknown length
59042: 60341: contig of 1299 bp in length
60341: 60441: gap of unknown length
60441: 61693: contig of 1252 bp in length
61693: 61793: gap of unknown length
61793: 63315: contig of 1522 bp in length
63315: 63415: gap of unknown length
63415: 64571: contig of 1156 bp in length
64571: 64671: gap of unknown length
64671: 66083: contig of 1412 bp in length
66083: 66183: gap of unknown length
66183: 67516: contig of 1333 bp in length
67516:

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2002, 03:28:26 ; Search time 2214 Seconds
(without alignments)
2567.577 Million cell updates/sec

Title: US09897438BK-2

Perfect score: 351

Sequence: 1 GARCARTGYGNGACNATNAT.....AYGARGCNTGYTGGCNCVTN 351

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_othr:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	12	3.4	321	17 A0226184	A0226184 HS.2022_B
2	12	3.4	377	14 C97776	C97776 C97776 Rice
3	12	3.4	549	17 A0674732	A0674732 HS.5485_B
c 4	12	3.4	831	12 BE705775	BE705775 Sc02_04b1
5	12	3.4	1714	17 AG029608	AG029608 Pan trogl
c 6	11	3.1	202	14 W20272	W20272 zb43h08.r1

7	11	3.1	226	14	N65130
8	11	3.1	333	14	C74396
c 9	11	3.1	361	13	BJ073290
10	11	3.1	381	14	F23596
c 11	11	3.1	408	9	AL045493
12	11	3.1	437	9	AI997638
13	11	3.1	530	9	AI562717
c 14	11	3.1	604	17	FR0043054
15	11	3.1	613	13	BJ503599
16	11	3.1	634	13	BI959766
17	11	3.1	1051	17	CNS008E4
c 18	11	3.1	1068	17	CNS05D8A
c 19	11	3.1	1071	13	BG985678
20	11	3.1	1533	17	AG107622
c 21	10	2.8	145	10	AV849260
c 22	10	2.8	148	14	T49675
c 23	10	2.8	190	13	BJ467155
c 24	10	2.8	193	14	T63392
c 25	10	2.8	222	17	B04666
c 26	10	2.8	253	9	AU210723
c 27	10	2.8	254	14	T02496
c 28	10	2.8	257	10	BB157379
c 29	10	2.8	267	13	BJ438963
30	10	2.8	278	13	BJ368840
c 31	10	2.8	283	13	BJ377516
32	10	2.8	300	9	AV199737
33	10	2.8	300	13	BJ135481
c 34	10	2.8	315	14	F08819
c 35	10	2.8	316	14	W03213
c 36	10	2.8	316	14	W03213
37	10	2.8	317	13	BJ366180
c 38	10	2.8	322	13	BJ399800
c 39	10	2.8	323	9	AA681046
c 40	10	2.8	324	14	D32970
c 41	10	2.8	378	14	T34001
c 42	10	2.8	387	12	BG520970
c 43	10	2.8	389	17	FR0028242
c 44	10	2.8	393	13	BJ131990
c 45	10	2.8	400	9	AL360472

ALIGNMENTS

RESULT 1
A0226184/c
LOCUS
DEFINITION
HS_2022_B1_E12_T7 CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2022 Col=23 Row=J, DNA sequence.
ACCESSION
A0226184
VERSION
A0226184.1
KEYWORDS
GSS.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 321)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
JOURNAL
MEDLINE
COMMENT
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallaceu.washington.edu
Sequence Tagged Connector
Plate: 2022 Row: J Column: 23

```

Class: BAC ends
High quality sequence stop: 321.
Location/Qualifiers
1..321
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=2022 Col=23 Row=J"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/notes="Organ: sperm; Vector: pBelobAC11; BAC Clones in
E-Coli DH10B"
93 a 63 c 56 g 102 t 7 others
BASE COUNT 93 a 63 c 56 g 102 t 7 others
ORIGIN

Query Match
Best Local Similarity 3.4%; Score 12; DB 17; Length 321;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 NACNATNGCA 23
Db 247 NACNATNGCA 236

RESULT 2
C97776
LOCUS
DEFINITION C97776 Rice callus Oryza sativa (japonica cultivar-group) cDNA
ACCESSION clone C62743_12, mRNA sequence.
VERSION C97776.1 GI:3760522
KEYWORDS EST.
ORGANISM Oryza sativa (japonica cultivar-group).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 377)
Sasaki, T. and Minobe, Y.
Rice cDNA from callus
Unpublished (1994)
Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@abrr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/
PROJECT = "RGP"
POLYA=NO.
Location/Qualifiers
1..377
/organism="Oryza sativa (japonica cultivar-group)"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="C62743_12"
/clone_lib="Rice callus"
/notes="Vector: pBluescript II SK+; Site_1: SalI; Site_2:
NotI; cDNA prepared from rice callus mRNAs by using
oligo(dT) as a primer and ligating to the SalI-NotI site
of pBluescript II SK+ phagemid."
95 a 76 c 83 g 110 t 13 others
BASE COUNT 95 a 76 c 83 g 110 t 13 others
ORIGIN

Query Match
Best Local Similarity 3.4%; Score 12; DB 14; Length 377;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 ACNACNACNGCN 99
Db 173 ACNACNACNGCN 184

RESULT 3

```

```

A0674732
LOCUS
DEFINITION HS_5485_B1_G05_T7A RPCI-11 Human Male BAC Library GSS 24-JUN-1999
genomic clone Plate=1061 Col=9 Row=N, DNA sequence.
ACCESSION A0674732
VERSION A0674732.1 GI:5207478
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 549)
Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
Hood, L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
99380589
Contact: Mahairas GG, Wallace JC, Hood L
High throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 1061 row: N column: 9
Seq primer: T7
Class: BAC ends
High quality sequence stop: 549.
Location/Qualifiers
1..549
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=1061 Col=9 Row=N"
/clone_lib="RPCI-11 Human Male BAC Library"
/sex="male"
/notes="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACe3.6 vector at EcoRI sites"
158 a 112 c 111 g 159 t 9 others
BASE COUNT 158 a 112 c 111 g 159 t 9 others
ORIGIN

Query Match
Best Local Similarity 3.4%; Score 12; DB 17; Length 549;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 NACNACNACNTG 80
Db 28 NACNACNACNTG 39

RESULT 4
BE705775/c
LOCUS
DEFINITION BE705775 SC02_04b10_A sc02_AAFc_EC0RC_cold_stressed_winter_rye_seedlings
Scalae cereale cDNA clone sc02_04b10, mRNA sequence.
ACCESSION BE705775
VERSION BE705775.1 GI:10094040
KEYWORDS EST.
SOURCE rye.
ORGANISM Scalae cereale
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Secale.

```



```

REFERENCE
AUTHORS Singh,J.A., Piche,C., Couroux,P., De Moors,A., Harris,L.J., Hattori
        .J.I., Ouellet,T., Robert,L.S., Sprott,D. and Tinker,N.A.
TITLE Expressed Sequence Tags from Cold-Stressed Winter Rye Seedlings
JOURNAL Unpublished (2000)
COMMENT Contact: Singh,J.A.
        Eastern Cereal and Oilseed Research Centre
        Agriculture and Agri-food Canada
        KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A
        OC6, Canada
        Tel: (613) 759-1662
        Fax: (613) 759-1701
        Email: singhja@em.agr.ca.

FEATURES
source
1. .831
   /organism="Secale cereale"
   /cultivar="Puma (winter rye)"
   /db_xref="taxon:4550"
   /clone="Sc02.04b10"
   /clone_lib="Sc02_AAFc_ECORC_cold_stressed_winter_rye_seedl
   ings"
   /tissue_type="leaf, crown"
   /dev_stage="seedling three-leaf stage"
   /note="Vector: Bluescript SK+/XhoI-EcoRI; Site.1: Eco RI;
   Site.2: Xho I; Sampled three-leaf seedlings treated for
   one week at 20C, 12 hrs light/day. Library made with
   Stratagene UNIZAP XR Kit(not packaged). cDNA is directly
   ligated into SK+/XhoI-EcoRI, then electroporated into
   TOP10 cells (Invitrogen)."
BASE COUNT 78 a 180 c 211 g 264 t 98 others
ORIGIN

Query Match 3.4%; Score 12; DB 12; Length 831;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 TNACNACNACNT 79
   |||||
Db 705 TNACNACNACNT 694

RESULT 5
AG029608 1714 bp DNA linear GSS 01-NOV-2001
LOCUS Pan troglodytes DNA, clone: PTB-001L22.R, genomic survey sequence.
DEFINITION AG029608
ACCESSION AG029608.1 GI:16556481
VERSION GSS.
KEYWORDS Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male
SOURCE BAC Library clone:PTB-001L22.R.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
        Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE BAC end sequences of Library PTB
JOURNAL Unpublished
REFERENCE
AUTHORS 2 (bases 1 to 1714)
        Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
        Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
        and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
        1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
        (E-mail:chimbases@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
        Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end
        was generated during the R&D process and may have higher chance of
        clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY

```

```

Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI
Location/Qualifiers
1. .1714
   /organism="Pan troglodytes"
   /db_xref="taxon:9598"
   /clone="PTB-001L22.R"
   /sex="male"
   /cell_type="lymphoblast"
   /clone_lib="PTB Chimpanzee Male BAC Library"
BASE COUNT 527 a 297 c 472 g 325 t 93 others
ORIGIN

Query Match 3.4%; Score 12; DB 17; Length 1714;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 TNACNACNACNT 79
   |||||
Db 1616 TNACNACNACNT 1627

RESULT 6
W20272/c 202 bp mRNA linear EST 03-MAY-1996
LOCUS zb43808.rl Soares_fetal_lung_NbHL19W Homo sapiens cDNA clone
DEFINITION IMAGE:306399 5' similar to WP:T09A5.6 CE01087 ; mRNA sequence.
ACCESSION W20272
VERSION W20272.1 GI:1296206
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
        .M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
        Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
        .R., Williamson,A., Wohlmann,P. and Wilson,R.
        The WashU-Merck EST Project
        Contact: Wilson RK
        Washington University School of Medicine
        4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
        Tel: 314 286 1800
        Fax: 314 286 1810
        Email: est@watson.wustl.edu
        This clone is available royalty-free through LLNL ; contact the
        IMAGE Consortium (info@image.llnl.gov) for further information.
        Trace considered overall poor quality
        Possible reversed clone: similarity on wrong strand
        Seq primer: mob.REGA+ET
        High quality sequence stop: 1.
FEATURES
source
1. .202
   /organism="Homo sapiens"
   /db_xref="GDB:124981"
   /db_xref="taxon:9606"
   /clone="IMAGE:306399"
   /clone_lib="Soares_fetal_lung_NbHL19w"
   /dev_stage="19 weeks"
   /lab_host="DH10B (ampicillin resistant)"
   /note="Organ: lung; Vector: pT7T3D (Pharmacia) with a
   modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st
   strand cDNA was primed with a Not I - oligo(dT) primer
   [5'-TCTTACCAATCTGAAGTGGAGCGCCGAATTTTTTTTTTTT-3'],
   double-stranded cDNA was size selected, ligated to Eco RI
   adapters (Pharmacia), digested with Not I and cloned into
   the Not I and Eco RI sites of a modified pT7T3 vector
   (Pharmacia). Library went through one round of
   normalization to a Cot = 5. Library constructed by Bento
   Soares and M.Fatima Bonaldo. This library was constructed

```

```

from the same fetus as the fetal heart library, Soares
fetal heart NbHH19W."
BASE COUNT      58 a      32 c      49 g      50 t      13 others
ORIGIN

Query Match      3.1%; Score 11; DB 14; Length 202;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 NACNATNATGC 22
|||||
Db 52 NACNATNATGC 42

RESULT 7
N65130
LOCUS
DEFINITION      226 bp mRNA linear EST 05-JAN-1998
ACCESSION      N65130
VERSION
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 226)
Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh
,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M., Retzel
,E. and Somerville,C.
Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
95148729
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcn@ibm.cl.msu.edu
Seq primer: T7 dye primer.
FEATURES
source
1..226
/organism="Arabidopsis thaliana"
/strain="var columbia"
/db_xref="taxon:3702"
/clone="228D21T7"
/clone_lib="Lambda-PRL2"
/notes="Vector: lambda Zip-Lox; Site_1: Sal; Site_2: Not;
Lambda PRL2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
silicles. The vector is BRL's lambda Zip-Lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dt primed cDNA."
BASE COUNT      77 a      41 c      49 g      50 t      9 others
ORIGIN

Query Match      3.1%; Score 11; DB 14; Length 226;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 TNACNACNACN 78
|||||
Db 54 TNACNACNACN 54

from the same fetus as the fetal heart library, Soares
fetal heart NbHH19W."
BASE COUNT      58 a      32 c      49 g      50 t      13 others
ORIGIN

Query Match      3.1%; Score 11; DB 14; Length 202;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 NACNATNATGC 22
|||||
Db 52 NACNATNATGC 42

RESULT 7
N65130
LOCUS
DEFINITION      226 bp mRNA linear EST 05-JAN-1998
ACCESSION      N65130
VERSION
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 226)
Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh
,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M., Retzel
,E. and Somerville,C.
Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)
95148729
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcn@ibm.cl.msu.edu
Seq primer: T7 dye primer.
FEATURES
source
1..226
/organism="Arabidopsis thaliana"
/strain="var columbia"
/db_xref="taxon:3702"
/clone="228D21T7"
/clone_lib="Lambda-PRL2"
/notes="Vector: lambda Zip-Lox; Site_1: Sal; Site_2: Not;
Lambda PRL2 is a cDNA library derived from equal
quantities of 4 pools of mRNA. The mRNA sources were 1) 7
day germinated etiolated seedlings; 2) tissue culture
grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
same plants as 3 but aerial tissue (stems, flowers and
silicles. The vector is BRL's lambda Zip-Lox. The cDNA
inserts were directionally cloned with Sal-Not arms using
oligo dt primed cDNA."
BASE COUNT      77 a      41 c      49 g      50 t      9 others
ORIGIN

Query Match      3.1%; Score 11; DB 14; Length 226;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 TNACNACNACN 78
|||||
Db 54 TNACNACNACN 54

```

```

RESULT 8
C74396
LOCUS
DEFINITION      333 bp mRNA linear EST 04-APR-2002
ACCESSION      C74396
VERSION
KEYWORDS
SOURCE
ORGANISM
Oryza sativa (japonica cultivar-group).
Oryza sativa (japonica cultivar-group).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 333)
Sasaki,T. and Yamamoto,K.
Rice cDNA from panicle
Unpublished (1997)
JOURNAL
COMMENT
Contact: Takuji Sasaki
National Institute of Agrobiological Resources
Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
305-8602, Japan
Tel: 81-298-38-7441
Fax: 81-298-38-7468
Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
PROJECT = RGP.
FEATURES
source
1..333
/organism="Oryza sativa (japonica cultivar-group)"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="E31045.1A"
/dev_stage="shorter than 3cm"
/note="Organ: panicle"
BASE COUNT      41 a      143 c      87 g      46 t      16 others
ORIGIN

Query Match      3.1%; Score 11; DB 14; Length 333;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 ACNACNACNCG 98
|||||
Db 186 ACNACNACNCG 196

RESULT 9
BJ073290/c
LOCUS
DEFINITION      361 bp mRNA linear EST 11-DEC-2001
ACCESSION      BJ073290
VERSION
KEYWORDS
SOURCE
ORGANISM
Xenopus laevis
African clawed frog.
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 361)
Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-I,T. and Kohara
,Y.
Expressed genes in X. laevis embryo
Unpublished (2001)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES
source
1..361

```

```

/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="XL109i15"
/clone_lib="NIBB Mochii normalized Xenopus tailbud
library"
/tissue_type="whole embryo"
/dev_stage="stage 25"
/notes="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
were oligo-dT primed and directionally cloned. Staging
according to Nieuwkoop and Faber. Library is subtracted
and was constructed by N. Garrett and A.M. Zorn,
(Wellcome/CRC Institute)."
BASE COUNT      74 a    62 c    89 g    121 t    15 others
ORIGIN
Query Match      3.1%; Score 11; DB 13; Length 361;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 TNACNACNACN 78
|||||
Db 349 TNACNACNACN 339

RESULT 10
LOCUS      F23596      381 bp mRNA linear EST 13-MAY-1999
DEFINITION HSP004953 HM3 Homo sapiens cDNA clone NOTAVAIL04953, mRNA sequence.
ACCESSION F23596
VERSION F23596.1 GI:4809604
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 381)
AUTHORS Lanfranchi,G., Muraro,T., Caldara,F., Pacchioni,B., Pallavicini,A.,
Pandolfo,D., Toppo,S., Trevisan,S., Scarso,S. and Valle,G.
IDENTIFICATION OF 4370 EXPRESSED SEQUENCE TAGS FROM A
3'-END-SPECIFIC cDNA LIBRARY OF HUMAN SKELETAL MUSCLE BY DNA
SEQUENCING AND FILTER HYBRIDIZATION
Genome Res. 6 (1), 35-42 (1996)
MEDLINE 96276048
CONTACT: Valle G.
CRIBI Biotechnology Centre
University of Padua
Via Trieste 75, 35121 Padua, Italy
ABI Chromatograms and other information are available on WWW at
http://group.bio.unipd.it.
FEATURES
Location/Qualifiers
1..381
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="NOTAVAIL04953"
/clone_lib="HM3"
/sex="female"
/tissue_type="pectoral muscle (after mastectomy)"
/notes="Vector: pcDNAII (Invitrogen); Site_1: BstXI;
Site_2: NotI; The library was constructed by G.
Lanfranchi. This library is not subtracted nor normalized.
The first strand cDNA was primed with a biotinylated
oligo-dT-NotI primer
(5'-biotin-AACCGGCTCGAGCGCGCTTTTCTTTTCTTTT-3'). The
ds cDNA was sonicated and size-selected in the range
350-550 bp. The 3' specific fragments were selected by
streptavidin coated magnetic beads, ligated to
non-palindromic BstXI adapters, NotI digested and
directionally cloned into BstXI-NotI cut pcDNAII vector."
BASE COUNT      91 a    98 c   110 g    68 t   14 others
ORIGIN
Query Match      3.1%; Score 11; DB 14; Length 381;
Best Local Similarity 100.0%; Pred. No. 26;

```

```

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 ACNACNACNCG 98
|||||
Db 197 ACNACNACNCG 207

RESULT 11
LOCUS      AL045493      408 bp mRNA linear EST 09-JUL-1999
DEFINITION DKFZp434i1165_r1 434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION AL045493
VERSION AL045493
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 408)
AUTHORS Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and Wiemann
,J.S.
EST (Duesterhoeft, et al.)
Unpublished (1999)
CONTACT: Duesterhoeft A
MIPS
Am Klopferstritz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); sequenced by Qiagen within the cDNA
sequencing consortium of the German Genome Project. s1 sequence
also available.
This clone is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
Location/Qualifiers
1..408
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp434i1165"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/notes="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT      82 a    87 c    74 g   120 t   45 others
ORIGIN
Query Match      3.1%; Score 11; DB 9; Length 408;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 GCNGTNACNNTT 41
|||||
Db 109 GCNGTNACNNTT 99

RESULT 12
LOCUS      AI997638      437 bp mRNA linear EST 08-SEP-1999
DEFINITION thaliana cDNA clone 701669602, mRNA sequence.
ACCESSION AI997638
VERSION AI997638.1 GI:5844543
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 437)
AUTHORS Chen,J., Momiyama,M., Chan,E., Mooney,M., Carroon,B., Gilliland,D.,
Wang,X., Hillman,J., Guegler,K., Kim,C., Doyle,M., Brzoska,P.,

```

Gorgone,G., Burns,D., Griffin,J., Mouanoutoua,M., Nguyen,D., Tan,R., Rose,M., Warren,B., Ton,B., Kastury,K., Borillo,C., Carpio,T., Policky,J., Suzuki,G., Argentine,C., Shah,S., Nobrega,A., Murry,L., Turner,C., Krikorian,S., Elder,L. and Hanson,D.
 Arabidopsis thaliana Gene Expression MicroArray
 Unpublished (1999)
 Contact: David Smoller, Ph.D.
 Genome Systems, Inc., a wholly owned subsidiary of Incyte Pharmaceuticals, Inc.
 4633 World Parkway Circle, St. Louis, MO 63134, USA
 Tel: 877-577-2733
 Fax: 314-427-3324
 Email: service@genomesystems.com.

FEATURES

source

1. 437
 Location/Qualifiers
 /organism="Arabidopsis thaliana"
 /cultivar="Columbia Col-0"
 /db_xref="taxon:3702"
 /clone="701669602"
 /clone_lib="A. thaliana, Columbia Col-0, rosette-1"
 /tissue_type="rosette"
 /dev_stage="4 - 7 weeks"

/note="Vector: pSPORT; Site_1: NotI; Site_2: SalI; cDNA library was derived from untreated rosette tissue from Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks. Plants were grown in 1:1:1 peat moss/vermiculite/perlite soil at 22 deg. C +/- 3 deg. C under constant light, and watered with fertilizer. cDNA synthesis was initiated using a NotI-oligo(dT) primer. Double-stranded cDNA was blunted, ligated to SalI adaptors, digested with NotI, size-selected, and cloned into the NotI and SalI sites of the pSPORT vector."

BASE COUNT 155 a 87 c 72 g 88 t 35 others

ORIGIN

Query Match 3.1%; Score 11; DB 9; Length 437;
 Best Local Similarity 100.0%; Pred. No. 27;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 238 ACNGTNATNCA 248

Db 54 ACNGTNATNCA 64

RESULT 13

AI562717

LOCUS

DEFINITION TENS2704 T. cruzi epimastigote normalized cDNA Library Trypanosoma cruzi cDNA clone 2704 5', mRNA sequence.

ACCESSION AI562717.1 GI:4514062

VERSION

KEYWORDS

SOURCE

ORGANISM Trypanosoma cruzi.

Trypanosoma cruzi

Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;

Trypanosoma; Schizotrypanum.

1 (bases 1 to 530)

Verdun,R.E., Di Paolo,N.C., Urmenyi,T.P., Rondinelli,E., Frasch

,A.C.C. and Sanchez,D.O.

Gene discovery through expressed sequence tag sequencing in

trypanosoma cruzi

Infect. Immun. 66 (11), 5393-5398 (1998)

99003155

Contact: Sanchez D.O.

Instituto de Investigaciones Biotecnologicas (Univ. Nac. de Gral

San Martin)

Av. Gral Paz entre Albarillos y Constituyentes, INTI edificio 24

cp(4650) San Martin, Prov. de Bs As. Argentina

Tel: (54-1)752-9639 or (54-1)752-0021

Fax: (54-1)752-0021 or (54-1)752-9639

Email: dsanchez@inti.gov.ar

Seq primer: T7.

FEATURES

Location/Qualifiers

1. 437

Location/Qualifiers

/organism="Trypanosoma cruzi"

/strain="Cl-Brenner"

/db_xref="taxon:5693"

/clone="2704"

/clone_lib="T. cruzi epimastigote normalized cDNA Library"

/cell_type="epimastigote"

/note="cDNA library constructed with oligo dT primed epimastigote mRNA and cloned in pC7t318D phagemid with modified polylinker (PHARMACIA)"

BASE COUNT 110 a 143 c 131 g 115 t 31 others

ORIGIN

Query Match 3.1%; Score 11; DB 9; Length 530;
 Best Local Similarity 100.0%; Pred. No. 28;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 69 NACNACNACNT 79

Db 394 NACNACNACNT 404

RESULT 14

FR0043054/c

LOCUS

DEFINITION Fugu rubripes GSS sequence, clone 158G14aC4, genomic survey

sequence.

ACCESSION AL130546.1 GI:6112492

VERSION

KEYWORDS GSS; genome survey sequence.

SOURCE Takifugu rubripes

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes;

Tetraodontidae; Takifugu.

1 (bases 1 to 604)

Elgar,G., Clark,M.S., Smith,S., Meek,S., Warner,S., Edwards,Y.J.K.,

Umrani,Y., Williams,G. and Brenner,S.

Direct Submission

Submitted (11-OCT-1999) MRC Human Genome Mapping Project Resource

Centre, Hinxton, Cambridge, CB10 1SB. UK Email:

biohelp@hmp.mrc.ac.uk

Vector: pBluescript II KS

V-type: phagemid

PRIMER: KS

DESCR:

One pass dye-terminator sequencing of cosmid cloned genomic

sequence.

FEATURES

source

1. 604

Location/Qualifiers

/organism="Takifugu rubripes"

/db_xref="taxon:31033"

/clone="158G14aC4"

/clone_lib="cosmid 158G14"

BASE COUNT 134 a 115 c 103 g 170 t 82 others

ORIGIN

Query Match 3.1%; Score 11; DB 17; Length 604;
 Best Local Similarity 100.0%; Pred. No. 30;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 CNACNACNCGCN 99

Db 341 CNACNACNCGCN 331

RESULT 15

BJ503599

LOCUS

DEFINITION BJ503599 MF01FSA cDNA Oryzias latipes CDNA clone MF01FSA004A17 3',

ORIGIN

Query Match 3.1%; Score 11; DB 17; Length 604;
 Best Local Similarity 100.0%; Pred. No. 30;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 CNACNACNCGCN 99

Db 341 CNACNACNCGCN 331

RESULT 15

BJ503599

LOCUS

DEFINITION BJ503599 MF01FSA cDNA Oryzias latipes CDNA clone MF01FSA004A17 3',

ORIGIN

Query Match 3.1%; Score 11; DB 17; Length 604;
 Best Local Similarity 100.0%; Pred. No. 30;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 CNACNACNCGCN 99

Db 341 CNACNACNCGCN 331

RESULT 15

BJ503599

LOCUS

DEFINITION BJ503599 MF01FSA cDNA Oryzias latipes CDNA clone MF01FSA004A17 3',

ORIGIN

Query Match 3.1%; Score 11; DB 17; Length 604;
 Best Local Similarity 100.0%; Pred. No. 30;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 CNACNACNCGCN 99

Db 341 CNACNACNCGCN 331

RESULT 15

BJ503599

LOCUS

DEFINITION BJ503599 MF01FSA cDNA Oryzias latipes CDNA clone MF01FSA004A17 3',

ORIGIN

Query Match 3.1%; Score 11; DB 17; Length 604;
 Best Local Similarity 100.0%; Pred. No. 30;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 CNACNACNCGCN 99

Db 341 CNACNACNCGCN 331

RESULT 15

BJ503599

LOCUS

DEFINITION BJ503599 MF01FSA cDNA Oryzias latipes CDNA clone MF01FSA004A17 3',

ORIGIN

Query Match 3.1%; Score 11; DB 17; Length 604;
 Best Local Similarity 100.0%; Pred. No. 30;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 CNACNACNCGCN 99

Db 341 CNACNACNCGCN 331

RESULT 15

BJ503599

LOCUS

DEFINITION BJ503599 MF01FSA cDNA Oryzias latipes CDNA clone MF01FSA004A17 3',

ORIGIN

Query Match 3.1%; Score 11; DB 17; Length 604;
 Best Local Similarity 100.0%; Pred. No. 30;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 CNACNACNCGCN 99

Db 341 CNACNACNCGCN 331

RESULT 15

BJ503599

LOCUS

DEFINITION BJ503599 MF01FSA cDNA Oryzias latipes CDNA clone MF01FSA004A17 3',

ORIGIN

Query Match 3.1%; Score 11; DB 17; Length 604;
 Best Local Similarity 100.0%; Pred. No. 30;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 CNACNACNCGCN 99

Db 341 CNACNACNCGCN 331

RESULT 15

BJ503599

LOCUS

DEFINITION BJ503599 MF01FSA cDNA Oryzias latipes CDNA clone MF01FSA004A17 3',

ORIGIN

Query Match 3.1%; Score 11; DB 17; Length 604;
 Best Local Similarity 100.0%; Pred. No. 30;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 CNACNACNCGCN 99

Db 341 CNACNACNCGCN 331

RESULT 15

BJ503599

LOCUS

DEFINITION BJ503599 MF01FSA cDNA Oryzias latipes CDNA clone MF01FSA004A17 3',

ORIGIN

Query Match 3.1%; Score 11; DB 17; Length 604;
 Best Local Similarity 100.0%; Pred. No. 30;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 CNACNACNCGCN 99

Db 341 CNACNACNCGCN 331

RESULT 15

BJ503599

LOCUS

DEFINITION BJ503599 MF01FSA cDNA Oryzias latipes CDNA clone MF01FSA004A17 3',

ORIGIN

Query Match 3.1%; Score 11; DB 17; Length 604;
 Best Local Similarity 100.0%; Pred. No. 30;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 CNACNACNCGCN 99

Db 341 CNACNACNCGCN 331

RESULT 15

BJ503599

LOCUS

DEFINITION BJ503599 MF01FSA cDNA Oryzias latipes CDNA clone MF01FSA004A17 3',

ORIGIN

Query Match 3.1%; Score 11; DB 17; Length 604;
 Best Local Similarity 100.0%; Pred. No. 30;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 CNACNACNCGCN 99

Db 341 CNACNACNCGCN 331

RESULT 15

BJ503599

LOCUS

DEFINITION BJ503599 MF01FSA cDNA Oryzias latipes CDNA clone MF01FSA004A17 3',

ORIGIN

Query Match 3.1%; Score 11; DB 17; Length 604;
 Best Local Similarity 100.0%; Pred. No. 30;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 CNACNACNCGCN 99

Db 341 CNACNACNCGCN 331

RESULT 15

BJ503599

LOCUS

DEFINITION BJ503599 MF01FSA cDNA Oryzias latipes CDNA clone MF01FSA004A17 3',

ORIGIN

Query Match 3.1%; Score 11; DB 17; Length 604;
 Best Local Similarity 100.0%; Pred. No. 30;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 CNACNACNCGCN 99

Db 341 CNACNACNCGCN 331

RESULT 15

BJ503599

LOCUS

DEFINITION BJ503599 MF01FSA cDNA Oryzias latipes CDNA clone MF01FSA004A17 3',

ORIGIN

Query Match 3.1%; Score 11; DB 17; Length 604;
 Best Local Similarity 100.0%; Pred. No. 30;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 CNACNACNCGCN 99

Db 341 CNACNACNCGCN 331

RESULT 15

```

VERSION BJ503599.1 GI:22155561
KEYWORDS EST.
SOURCE Japanese medaka.
ORGANISM Oryzias latipes

REFERENCE 1 (bases 1 to 613)
AUTHORS Kohara,Y., Shin-i,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
TITLE Medaka EST Project in Takeda's lab
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES             source
     1..613
        /organism="Oryzias latipes"
        /strain="d-rf"
        /db_xref="taxon:8090"
        /clone="MF01FSA004A17"
        /clone_lib="MF01FSA cDNA"
        /sex="mixture of female and male"
        /tissue_type="whole embryo"
        /dev_stage="fry stage 40"

BASE COUNT 159 a 123 c 105 g 206 t 20 others
ORIGIN

Query Match 3.1%; Score 11; DB 13; Length 613;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 ACNATNATGCA 23
   |
Db 279 ACNATNATGCA 289

```

Search completed: November 7, 2002, 04:58:37
 Job time : 2222 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run 'on: November 6, 2002, 23:35:35 ; Search time 301 Seconds
(without alignments)
2626.084 Million cell updates/sec

Title: US09897438BK-2

Perfect score: 351

Sequence: 1 GARCARTGYGACNATNAT.....AYGARGCNTGYTGGCNCVTN 351

Scoring table: IDENTITY_NUC

Gap 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*

2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*

3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*

4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*

5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*

6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*

7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*

8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*

9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*

10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*

11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*

12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*

13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*

14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*

15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*

16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*

17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*

18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*

19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*

20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*

21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*

22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*

23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*

24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	249.4	71.1	351	ABL40165	Mouse reelin prote
2	249.4	71.1	2745	AAD22754	Mus musculus trunc
3	249.4	71.1	11673	ABA92603	Mouse reelin encod
4	249.4	71.1	11673	ABI99284	Mouse ischaemic co
5	233.8	66.6	11580	ABN96939	Gene #3437 used to
6	233.8	66.6	11580	ABA92604	Human reelin encod
7	233.8	66.6	11632	AA589484	DNA encoding novel
8	193.8	55.2	2025	AAD22779	Xenopus laevis int
9	193.8	55.2	2274	AAD22753	Xenopus laevis tru

C 10	53.4	15.2	413	22	AAK07720	Human brain expres
C 11	53.4	15.2	413	22	AAK33553	Human bone marrow
C 12	53.4	15.2	413	24	ABS08422	Human genome-deriv
C 13	44.8	12.8	87	22	AAK20384	Human brain expres
C 14	44.8	12.8	87	22	AAK46490	Human bone marrow
C 15	44.8	12.8	87	24	ABS20836	Human genome-deriv
C 16	38.4	10.9	4590	22	AAH24065	Yeast AOD904-asso
C 17	37.8	10.8	6741	21	AAAI0595	Gene encoding a su
C 18	36.2	10.3	2949	21	AAD00024	Degenerate DNA enc
C 19	36.2	10.3	2949	24	AAK37455	Human zfstaz2 degen
C 20	36	10.3	759	24	AAD38781	Mouse zlm24 dege
C 21	36	10.3	14041	22	AAH48024	Internal control B
C 22	36	10.3	14041	22	AAH48024	Internal control B
C 23	35.8	10.2	10732	21	AAAI0594	Gene encoding a su
C 24	35.6	10.1	371	22	AAI91689	Human polynucleoti
C 25	35.6	10.1	117574	24	AAI45288	Human KCNB1 gene.
C 26	35.4	10.1	485	22	ABA58007	Human foetal liver
C 27	35.4	10.1	485	22	AAK06081	Human brain expres
C 28	35.4	10.1	485	22	AAK31731	Human bone marrow
C 29	35.4	10.1	485	22	AAI37607	Probe #6293 used t
C 30	35.4	10.1	485	24	ABS06487	Human genome-deriv
C 31	35.4	10.1	1075	20	AAK84335	Stealth virus nucl
C 32	35.2	10.0	467	22	AAK62369	Human immune/haema
C 33	35.2	10.0	696	24	AAD22922	Human soluble IL-2
C 34	35.2	10.0	984	16	AAQ94336	Degenerate Alterom
C 35	35.2	10.0	3630	23	ABL27020	Drosophila melanog
C 36	35	10.0	7737	23	ABL02204	Human interleukin-
C 37	34.8	9.9	660	19	AAV31481	Human interleukin
C 38	34.8	9.9	660	20	AAZ40421	S. epidermidis ope
C 39	34.8	9.9	915	22	AAH53398	S. epidermidis gen
C 40	34.8	9.9	1210	22	AAH54976	Degenerate DNA of
C 41	34.8	9.9	1905	24	AAK98781	Degenerate DNA of
C 42	34.8	9.9	3446	22	AAH54934	S. epidermidis gen
C 43	34.6	9.9	263	14	AAQ61171	Human brain Expres
C 44	34.6	9.9	581	22	AAH43479	(R)-2-octanol deny
C 45	34.6	9.9	593	22	ABA61216	Human foetal liver

ALIGNMENTS

RESULT 1
ABL40165
ID ABL40165 standard; DNA; 351 BP.
AC ABL40165;
XX
XX

21-MAY-2002 (first entry)
Mouse reelin protein CR-50 epitope region encoding DNA SEQ ID NO:1.

Mouse; reelin protein CR-50 epitope region; elucidation; neuron;
cerebral disturbance; reelin protein; neuroprotective; gene; ds.

Mus musculus.

Key Location/Qualifiers
CDS 1..351
/*tag= a

FT /product= "reelin protein CR-50 epitope region"
FT /note= "no start or stop codons given"

PN JP2002017361-A.

XX 22-JAN-2002.

XX 04-JUL-2000; 2000JP-0202801.

XX 04-JUL-2000; 2000JP-0202801.

XX (RIKA) RIKAGAKU KENKYUSHO.

CC drugs or therapeutics for ischaemic diseases. AB199913 and AB199914
CC represent PCR primers for a mouse ischaemic condition related sequence,
CC which are used in the exemplification of the present invention.
XX
XX
SQ Sequence 11673 BP; 2831 A; 2985 C; 2985 G; 2872 T; 0 other;
Query Match 71.1%; Score 249.4; DB 24; Length 11673;
Best Local Similarity 57.7%; Pred. No. 2.3e-63;
Matches 202; Conservative 79; Mismatches 69; Indels 0; Gaps 0;
QY 1 GARCARTYGGNACNATNATGCAATGAGGNAAYGCNATNATNTTYTGCGARCCNTATGGNCCN 60
DB 970 GAGCAGTGTGGCACCACATCATGCGCAATGCTGCTCCTCTGTGAGCCGTACGCCCT 1029
QY 61 MNGARTYNACNACNACNTGYTTNAAVACNACNACNCGNWSNGNTNCAATTTYSNATN 120
DB 1030 CGAGCTGACACACATGCTGACACACACACAGATGTCCTCCAGTTTTCATT 1089
QY 121 GGNWSNGNWSNTGYMNTTYSNATYWSNGAYGCNWSNATNACNATNWSNATYGCNNAAR 180
DB 1090 GGGTCAGGATCATGCTGATTTAGTTACTCTGACCCAGCATCACTGTGTCATACGCCAAG 1149
QY 181 AAYAAACNCGNGAYTGGATNCARYTNGARAATNNGNCGNCCNWSNAAAYGTNWSNACN 240
DB 1150 AACAAATACCGCTGATGCTGATTCAGCTGGAGAAATAGAGCCCTTCCAATGTGAGCACA 1209
QY 241 GTNATNCAYATNTYNTAYTNCNGARGCNAARGNGARWSNGTNCARTTYCARTGG 300
DB 1210 GTATCCACATCTGTACCTCCCGAGAGAGCAGGAGGAGCGTGCGAGTTCCAGTTGG 1269
QY 301 AARCAGAYWSNTYMGNTNGNGARGTNTAYGARGCNTGYTGGGNYT 350
DB 1270 AACAGGACAGCCCTGCGAGTGGGTGAGGTGATGAGGCGCTGCTGGGCCCT 1319
RESULT 5
ABN96939
ID ABN96939 standard; DNA; 11580 BP.
XX
XX
AC ABN96939;
XX
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #3437 used to diagnose liver cancer.
XX
KW Gene: liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumor; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
XX
XX WO200229103-A2.
XX
XX 11-APR-2002.
XX
XX 02-OCT-2001; 2001WO-US30589.
XX
XX 02-OCT-2000; 2000US-237054P.
XX
XX (GENE-) GENE LOGIC INC.
XX
XX Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX WPI; 2002-426119/45.
XX
XX Diagnosing and detecting the progression of liver cancer,
PT hepatocellular carcinoma or metastatic liver tumor in a patient,
PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample -
XX
XX
PS Claim 1; SEQ ID NO 3437; 298pp; English.
XX
XX The invention relates to a novel method for diagnosing and detecting the

CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytotatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 11580 BP; 3014 A; 2696 C; 2753 G; 3116 T; 1 other;
Query Match 66.6%; Score 233.8; DB 24; Length 11580;
Best Local Similarity 56.0%; Pred. No. 1.1e-58;
Matches 196; Conservative 74; Mismatches 80; Indels 0; Gaps 0;
QY 1 GARCARTYGGNACNATNATGCAATGAGGNAAYGCNATNATNTTYTGCGARCCNTATGGNCCN 60
DB 860 GAACAGTGTGGCGCGATTTATGCAATGCGCAATGCCGTCACTTCTGTGAACCATATGGCCCA 919
QY 61 MNGARTYNACNACNACNTGYTTNAAVACNACNACNCGNWSNGNTNCAATTTYSNATN 120
DB 920 CGAAGCTGATACACAGCGCTTAATACACACACAGCTTCTGCTCCCAATTTTCATT 979
QY 121 GGNWSNGNWSNTGYMNTTYSNATYWSNGAYGCNWSNATNACNATNWSNATYGCNNAAR 180
DB 980 GGGTCAGGTTGATGCTGCTTTAGTTATTACAGCCCGAGCATCATCGTTATATATGCAAG 1039
QY 181 AAYAAACNCGNGAYTGGATNCARYTNGARAATNNGNCGNCCNWSNAAAYGTNWSNACN 240
DB 1040 AATAACTCTCGGACTGGATTCAGCTAGAGAAATAGAGCCCTTCCAATGTGAGCACA 1099
QY 241 GTNATNCAYATNTYNTAYTNCNGARGCNAARGNGARWSNGTNCARTTYCARTGG 300
DB 1100 ATCATCCATATCTCTACCTTCTCTGAGGAGCGCCAAAGGAGAGATGTCCCAATTCAGTGG 1159
QY 301 AARCAGAYWSNTYMGNTNGNGARGTNTAYGARGCNTGYTGGGNYT 350
DB 1160 AAGCAGGAAATCTCTGCTGAGGTGAGTGTATGAGGCGCTGCTGGGCCTT 1209
RESULT 6
ABA92604
ID ABA92604 standard; cDNA; 11580 BP.
XX
XX ABA92604;
XX
XX 21-MAR-2002 (first entry)
XX
XX Human reelin encoding cDNA SEQ ID NO:4.
DE
DE Human reelin; low density lipoprotein receptor; LDLR; neuroprotective;
KW extracellular glycoprotein; nootropic; antipapenic; Alzheimer's disease;
KW neurodegenerative disorder; neuronal regeneration; cognitive function;
KW lipid metabolism disease; memory; developmental disorder; gene; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 176..10558
FT /*tag= a
FT /product= "human reelin"
XX
XX US6323177-B1.
XX
XX 27-NOV-2001.
XX
XX 16-JUN-1999; 99US-0334220.

```

XX 16-JUN-1999; 9905-0334220.
PR (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
PA Curran T, D'Arcangelo G;
XX WPI; 2002-096596/13.
DR P-PSDB; ABB05007.
XX Novel composition useful for screening compounds that modulate Reelin
PT binding to low density lipoprotein receptor, comprising an isolated
PT Reelin polypeptide and low density lipoprotein receptor
XX Example 1; Column 75-84; 45pp; English.
XX The present invention describes a composition (I) comprising an
CC isolated reelin protein (II) bound to an isolated low density lipoprotein
CC receptor (LDR) (III). (II) is an extracellular glycoprotein of
CC approximately 385 kDa containing a small region of similarity with
CC R-spondin at the N terminus, a stretch of positively charged amino
CC acids at the C terminus, and a series of eight internal repeats of
CC 350-390 amino acids, each repeat containing two related sub-domains
CC that flank a pattern of conserved cysteine residues known as an
CC epidermal growth factor (EGF)-like motif. (I) has neuroprotective,
CC neurotropic and antilipemic activities, and can be used as a modulator
CC of reelin-LDR interaction. (I) is useful in screen for compounds that
CC modulate reelin binding to an LDR, in an assay system, where the assay
CC system comprises a microplate array and an automated robotic
CC microprocessor controlled system for adding and removing reagents to
CC the microplate array. The compounds identified by the above screening
CC method are useful as therapeutic agents to provide or alleviate a
CC diverse spectrum of diseases including neurodegenerative disorders such
CC as Alzheimer's disease, to facilitate neuronal regeneration after
CC injury, to prevent or alleviate lipid metabolism diseases, to enhance
CC cognitive functions and memory or to ameliorate other developmental
CC disorders. The present sequence encodes human reelin, which is used in
CC the exemplification of the present invention.
XX Sequence 11580 BP; 3014 A; 2696 C; 2753 G; 3116 T; 1 other;

Query Match 66.6%; Score 233.8; DB 24; Length 11580;
Best Local Similarity 56.0%; Pred. No. 1.1e-58;
Matches 196; Conservative 74; Mismatches 80; Indels 0; Gaps 0;

QY 1 GARCARTGYGNNACNATNATGCGAYGNAAYCGNCTNACNTTYTYTGARCCNTAYGGNCCN 60
DB 860 GAACAGTGTGGCGGATTTATGCGCAATGCCGTCCACCTCTGTGTGAACCATATGGCCCA 919
QY 61 MNGNARYTNACNACNACNTGYTTNAAVACNACNACNCSNNGNTYTNCAATTTWSNATN 120
DB 920 CGAGAACTGATTACCAAGCCCTTAATACAAACACAGCTTCTGTCTCCAAATTTCCATT 979
QY 121 GGNWSNGNNSNTGYMGNTTYWSNTAYWSNGAYCCNWSNATNACNCTNWSNTAYGCNNAAR 180
DB 980 GGGTCAGGTTTCATGTCGCTTAGTTATTCAGACCCAGCATCATCGTTATATGCGCAAG 1039
QY 181 AAYAAACNCGNAYTGGATNCAATNCAARAATNMGNCNCCNWSNAAAYTNWSNACN 240
DB 1040 AATAACTCTCGGACTGGATTCAGCTAGAGAAAATTAGAGCCCTTCCAAATGTCAGACA 1099
QY 241 GTNATNCAYATNTYNTAYTNCCNGARGARGCNAARGNGNWSNCTNCAATTYCARTGG 300
DB 1100 ATCATCATATCTCTACCTTCTGAGAGAGCCCAAGGGGAGAAATGTCCAATTTCACTGG 1159
QY 301 AACRAGAYNSNTNMGNGTNGNGARGTNTAYGARGCNTGYTGGGNYT 350
DB 1160 AAGCAGGAAAATCTTCGTGTAGGTGAAGTGTATGAGCCCTGCTGGGCTT 1209

RESULT 7
AAS89484
ID AAS89484 standard; cDNA; 11632 BP.

```

```

XX AAS89484;
XX 13-FEB-2002 (first entry)
XX DNA encoding novel human diagnostic protein #25288.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-0508631.
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG25297.
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX Claim 1; SEQ ID NO 25288; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human
XX diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 11632 BP; 3028 A; 2716 C; 2770 G; 3117 T; 1 other;

Query Match 66.6%; Score 233.8; DB 23; Length 11632;
Best Local Similarity 56.0%; Pred. No. 1.1e-58;
Matches 196; Conservative 74; Mismatches 80; Indels 0; Gaps 0;

QY 1 GARCARTGYGNNACNATNATGCGAYGNAAYCGNCTNACNTTYTYTGARCCNTAYGGNCCN 60
DB 860 GAACAGTGTGGCGGATTTATGCGCAATGCCGTCCACCTCTGTGTGAACCATATGGCCCA 919
QY 61 MNGNARYTNACNACNACNTGYTTNAAVACNACNACNCSNNGNTYTNCAATTTWSNATN 120
DB 920 CGAGAACTGATTACCAAGCCCTTAATACAAACACAGCTTCTGTCTCCAAATTTCCATT 979
QY 121 GGNWSNGNNSNTGYMGNTTYWSNTAYWSNGAYCCNWSNATNACNCTNWSNTAYGCNNAAR 180

```



```
DR WPI; 2002-019320/03.
DR P-PSDB; AAEL3605.
PT Novel truncated Reelin protein containing F-spondin domain and CR-50
PT recognition site of Reelin protein, but not having Reelin repeat site,
PT useful to treat diseases including agyria due to abnormal neuron
PT alignment
XX
XX Claim 8; Page 11-16; 47pp; English.
XX
XX The invention relates to a truncated Reelin protein comprising a
XX F-spondin domain and a CR-50 recognition site but no reelin protein
XX repeat site. Reelin is an essential molecule in developing a normal
XX laminated structure of cerebrum. The truncated reelin protein and its
XX DNA are useful for treating diseases including agyria, polymicrogyria,
XX and ectopic gray matter due to abnormal neuronal alignment. Truncated
XX reelin protein DNA is useful in gene therapy. The present sequence is
XX a cDNA encoding Xenopus laevis truncated reelin protein.
XX
XX Sequence 2274 BP; 660 A; 450 C; 469 G; 694 T; 1 other;
XX
XX Query Match 55.2%; Score 193.8; DB 24; Length 2274;
XX Best Local Similarity 50.6%; Pred. No. 3.1e-47;
XX Matches 177; Conservative 66; Mismatches 107; Indels 0; Gaps 0;
XX
XX QY 1 GARCARTGYGGNACNATNATGTCAYGGNAAYGCNCTNACNTTYTGYGARCNTAYGGNCCN 60
XX ||:||:||||| || |||||:|| || |||||:|||| ||:|||| ||
XX Db 847 GAGCAGTGTGGTGAATATGATGCTGGTGGCGAGTCAGCTTTTGTGATCCGTATGGACCA 906
XX :||:|||| || |||||:|| || |||||:|||| ||:|||| ||
XX QY 61 MNGARTYNACNACNACNTGYTYYNAAYACNACNACNWSNGTNYTNCARTTYWSNATN 120
XX :||:|||| || |||||:|| || |||||:|||| ||:|||| ||
XX Db 907 AGAATATGATACTGTTCAAATCAACACAACTACGGCAGCTGTTTGGCAGTTTCTATT 966
XX ||:|||| || |||||:|| || |||||:|||| ||:|||| ||
XX QY 121 GGWSNGNWSNTGYMGNYYWSNATYWSNGAYCCNWSNATNACNNGTWSNATYGCNNAAR 180
XX ||:|||| || |||||:|||| |||||:|||| ||:|||| ||
XX Db 967 GGGTCAGGATCGTGCAGGTCAGCTATTCAGACCCCTGGAATTTGTGGTCTACACAAAG 1026
XX ||:|||| || |||||:|||| |||||:|||| ||:|||| ||
XX QY 181 AAYAAACNGCNGAYTGGATNCARYTNGARAARATNMGNGCNCNWSNAAAYGTNWSNACN 240
XX ||:|||| || |||||:|||| |||||:|||| ||:|||| ||
XX Db 1027 AATPAACTACTAGTTGGATGCCATTTGAGAGAATATGTCCTCTCCATYGTATGACC 1086
XX ||:|||| || |||||:|||| |||||:|||| ||:|||| ||
XX QY 241 GTNATNCAYATNYTNTAYTNCNGARGCNAARGCNAARGNGARWSNGTNCARTTYCARTGG 300
XX ||:|||| || |||||:|||| |||||:|||| ||:|||| ||
XX Db 1087 ATCATTCACATTAATTTACCTACCTCCTGAAGCTAAAGAGAGAAATGTGAATTCGGTTGG 1146
XX ||:|||| || |||||:|||| |||||:|||| ||:|||| ||
XX QY 301 AARCARGAYSNYTNMGNGTNGNGARGTNTAYGARGCNTGYTGGGNYT 350
XX ||:|||| || |||||:|||| |||||:|||| ||:|||| ||
XX Db 1147 AGGCAGGAGAACATGCAGGCGAGGTGATGTGTATGAAGCCTGCTGGCACT 1196
XX ||:|||| || |||||:|||| |||||:|||| ||:|||| ||
XX
XX RESULT 10
XX AAK07720/c
XX ID AAK07720 standard; DNA; 413 BP.
XX
XX AAK07720;
XX
XX 05-NOV-2001 (first entry)
XX
XX Human brain expressed single exon probe SEQ ID NO: 7711.
XX
XX Human; brain expressed exon; gene expression analysis; probe;
XX microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX epilepsy; cancer; ss.
XX
XX Homo sapiens.
XX
XX WC200157275-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00667.
XX
XX 04-FEB-2000; 2000US-0180312.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX PI (MOLE-) MOLECULAR DYNAMICS INC.
```

```
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX PI WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX
XX Example 4; SEQ ID NO: 7711; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is one of the probes of the
XX invention.
XX
XX Sequence 413 BP; 134 A; 86 C; 67 G; 126 T; 0 other;
XX
XX Query Match 15.2%; Score 53.4; DB 22; Length 413;
XX Best Local Similarity 58.5%; Pred. No. 9.9e-06;
XX Matches 48; Conservative 14; Mismatches 20; Indels 0; Gaps 0;
XX
XX QY 1 GARCARTGYGGNACNATNATGTCAYGGNAAYGCNCTNACNTTYTGYGARCNTAYGGNCCN 60
XX ||:||||:|||| || |||||:|||| || || |||||:|||| ||:|||| ||
XX Db 323 GAACAGTGTGGCGGATATGATGCGCAATGCCGTACCTTCTGTGAACCATATGGCCCA 264
XX :||:|||| || |||||:|||| || || |||||:|||| ||:|||| ||
XX QY 61 MNGARTYNACNACNACNTGY 82
XX :||:|||| || |||||:|||| || || |||||:|||| ||:|||| ||
XX Db 263 CGAGAACTGGTAAGTATGTGCT 242
XX
XX RESULT 11
XX AAK33553/c
XX ID AAK33553 standard; DNA; 413 BP.
XX
XX AAK33553;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human bone marrow expressed single exon probe SEQ ID NO: 8110.
XX
XX Human; bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
XX Homo sapiens.
XX
XX WO200157276-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000668.
XX
XX 04-FEB-2000; 2000US-0180312.
XX
XX 26-MAY-2000; 2000US-0207456.
XX
XX 30-JUN-2000; 2000US-0608408.
XX
XX 03-AUG-2000; 2000US-0632366.
XX
XX 21-SEP-2000; 2000US-0234687.
XX
XX 27-SEP-2000; 2000US-0236359.
XX
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX PI
```

XX WPI; 2001-489900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
PS Example 4; SEQ ID NO: 8110; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention.
XX
SQ Sequence 413 BP; 134 A; 86 C; 67 G; 126 T; 0 other;

Query Match 15.2%; Score 53.4; DB 22; Length 413;
Best Local Similarity 58.5%; Pred. No. 9.9e-06;
Matches 48; Conservative 14; Mismatches 20; Indels 0; Gaps 0

OY 1 GARCARTGYGCNACNATWATCATGCGNAAYCNGCNAGTNTGTGGARCCNTA YGNCN 60
DB 323 GAACAGTGTCGCCGATTATGCATGGCAATGCCGTCACCCTTCTGTGAACCATATGGCCA 264
 |||:||||| | | |||||:| ||:|||| ||| ||| |||:|:| |||:|:| |||:|:| |||:
OY 61 MGNGARYTNACNACNACTGYG 82
DB 263 CGAGAACTGGTAAGTATGTGCT 242
 :| |:| | | | |:|:

RESULT 12
ABS08422/c
ID ABS08422 standard; DNA; 413 BP.

XX AC ABS08422;
XX DT 19-AUG-2002 (first entry)
XX DE Human genome-derived single exon probe from lung SEQ ID No 8413.
XX KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagenen syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
XX hyaline membrane disease.
OS Homo sapiens.
XX WO200186003-A2.
PN
XX
PD 15-NOV-2001.
XX
PP 30-JAN-2001; 2001WO-US000665.
XX
PR 04-FEB-2000; 2000US-180312P.
PR 26-MAY-2000; 2000US-207456P.
PR 30-JUN-2000; 2000US-060840B.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-234687P.
PR 27-SEP-2000; 2000US-236359P.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to

measure gene expression in human lung samples -

Claim 1; SEQ ID No 8413; 634pp; English.

The invention relates to a spatially-addressable set of single exon nucleic acid probes for measuring gene expression in a sample derived from human lung comprising single exon nucleic acid probes having one of 12614 nucleic acid sequences mentioned in the specification, or their complements or the 12387 open reading frames derived from the 12614 probes. Also included are a microarray comprising the novel set of probes; the novel set of probes which hybridise at high stringency to a nucleic acid expressed in the human lung; measuring gene expression in a sample derived from human lung, comprising (a) contacting the array with a collection of detectably labeled nucleic acids derived from human lung mRNA, and (b) measuring the label detectably bound to each probe of the array; identifying exons in a eukaryotic genome, comprising (a) algorithmically predicting at least one exon from genomic sequences of the eukaryote; and (b) detecting specific hybridisation of detectably labeled nucleic acids from eukaryote lung mRNA, to a single exon probe, having a fragment identical to the predicted exon, the probe is included in the above mentioned microarray; assigning exons to a single gene, comprising (a) identifying exons from genomic sequence by the method above and (b) measuring the expression of each of the exons in several tissues and/or cell types using hybridisation to a single exon microarrays having a probe with the exon, where a common pattern of expression of the exons in the tissues and/or cell types indicates that the exons should be assigned to a single gene; a peptide comprising one of 12011 sequences, mentioned in the specification, or encoded by the probes/open reading frames (ORF). The probes are used for gene expression analysis, and for identifying exons in a gene, particularly using human lung derived mRNA and for the study of lung diseases such as asthma, lung cancer, chronic obstructive pulmonary disease (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary haemoderiosis, pulmonary histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension and hyaline membrane disease. The present sequence is a single exon probe of the invention.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences.

Sequence 413 BP; 134 A; 86 G; 126 T; 0 other;

Query Match 15.2%; Score 53.4; DB 24; Length 413;
Best Local Similarity 58.5%; Pred. No. 9.9e-06;
Matches 48; Conservative 14; Mismatches 20; Indels 0; Gaps

QY 1 GARCARTGCGNACNATNGCCAYGGNAAGCGNTWACNTTGTGYGARCNTAATGGCCN 60
|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
Db 323 GAACAGTGTCGCCGATTATGCATGGCAATGCCCTCACCTCTGTGAACCATATGGCCCA 264

QY 61 MGNGARYTNACNACNCTGY 82
:||||:| | |||::
Db 263 CGAGAAGCTGGTAGTGTGCT 242

RESULT 13
AAK20384/C
ID AAK20384 standard; DNA; 87 BP.
AC AAK20384:
XX
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 20375.
XX
KW Human: brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer; ss.

XX	Homo sapiens.
OS	WO200157275-A2.
PN	09-AUG-2001.
PD	30-JAN-2001; 2001WO-US00667.
PF	04-FEB-2000; 2000US-0180312.
XX	26-MAY-2000; 2000US-0207456.
XX	30-JUN-2000; 2000US-0608408.
XX	03-AUG-2000; 2000US-0632366.
XX	21-SEP-2000; 2000US-0234687.
XX	27-SEP-2000; 2000US-0236359.
XX	04-OCT-2000; 2000GB-0024263.
PA	(MOLE-) MOLECULAR DYNAMICS INC.
PI	Penn SG, Hanzel DK, Chen W, Rank DR;
XX	WPI; 2001-483446/52.
XX	Single exon nucleic acid probes for analyzing gene expression in human brains -
PT	Example 4; SEQ ID NO: 20375; 650pp + Sequence Listing; English.
PS	The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human brain. They can be used to measure gene expression in brain cell samples, which may enable the diagnosis and improved treatment of nervous system diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia, epilepsy and cancers. The present sequence is one of the probes of the invention.
XX	Sequence 87 BP; 21 A; 25 C; 19 G; 22 T; 0 other;
SQ	Query Match 12.8%; Score 44.8; DB 22; Length 87; Best Local Similarity 67.8%; Pred. No. 0.0012; Matches 40; Conservative 9; Mismatches 10; Indels 0; Gaps 0
OY	1 GARCARTGYGGNACNATNATGCAYGGNAAYCGNCGTGACCCNTTGTGYGARCCNTA YGNCC 59 :: :: :: :: 59 GAACAGTGTCGCCGATTATCATGGCAATGCCGTCACCTTCTGTGAACCATATGSCC 1
Db	RESULT 14 AAK46490/c ID AAK46490 standard; DNA; 87 BP. XX AAK46490; XX AC XX DT XX DT XX DE XX DE XX Human bone marrow expressed single exon probe SEQ ID NO: 21047. XX Human; bone marrow expressed exon; gene expression analysis; probe; KW microarray; cancer; leukaemia; lymphoma; myeloma; ss. XX Homo sapiens. XX OS XX PN XX WO200157276-A2. XX PD XX 09-AUG-2001. XX 30-JAN-2001; 2001WO-US00668. XX PR XX 04-FEB-2000; 2000US-0180312. XX PR XX 26-MAY-2000; 2000US-0207456. XX PR XX 30-JUN-2000; 2000US-0608408. XX PR XX 03-AUG-2000; 2000US-0632366. XX PR XX 21-SEP-2000; 2000US-0234687.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2002, 00:47:00 ; Search time 68 seconds
(without alignments)
1582.993 Million cell updates/sec

Title: US09897438BK-2
Perfect score: 351
Sequence: 1 GARCARTGYGCGNACNATNAT.....AYGARGCNTGYTGGCGNYTN 351

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	249.4	71.1	11673	4	US-09-334-220-3
2	233.8	66.6	11580	4	US-09-334-220-4
3	39.2	11.2	7218	1	US-08-232-463-14
4	36.2	10.3	2949	4	US-09-412-554A-3
5	34.4	9.8	624	4	US-09-397-992A-3
6	34	9.7	37948	4	US-09-251-645-11
7	33.8	9.6	628	4	US-09-385-982-149
8	33.4	9.5	176373	3	US-09-128-155-17
9	33.2	9.5	1248	4	US-09-134-001C-2240
10	33.2	9.5	152331	3	US-09-128-155-16
11	33	9.4	957	4	US-09-134-001C-543
12	32.8	9.3	624	4	US-09-397-992A-6
13	32.6	9.3	2067	4	US-09-106-194-11
14	32.6	9.3	7498	2	US-08-816-693A-1
15	32.6	9.3	7498	3	US-08-885-291-1
16	32.6	9.3	7498	4	US-09-496-672-1
17	32.4	9.2	6743	3	US-08-932-280-1
18	32.4	9.2	12949	4	US-09-538-414-11
19	32.2	9.2	2294	4	US-09-643-597-123
20	32	9.1	5476	4	US-09-221-017B-147
21	31.8	9.1	31880	4	US-09-453-702B-242
22	31.8	9.1	246240	2	US-08-724-394A-20
23	31.8	9.1	246240	2	US-08-724-394A-21
24	31.8	9.1	246240	2	US-08-724-394A-22
25	31.6	9.0	361	1	US-08-053-131-171
26	31.6	9.0	361	1	US-08-096-762-171
27	31.6	9.0	2932	2	US-08-481-337A-5

28 31.6 9.0 2932 4 US-09-382-256-5 Sequence 5, Appli
29 31.6 9.0 2932 4 US-09-395-115-5 Sequence 5, Appli
30 31.6 9.0 2932 4 US-08-436-265-5 Sequence 5, Appli
31 31.6 9.0 2932 4 US-09-679-187-5 Sequence 5, Appli
32 31.6 9.0 2932 5 PCT-US95-05467-5 Sequence 5, Appli
33 31.4 8.9 2559 4 US-09-118-408-43 Sequence 43, Appli
34 31.4 8.9 2559 4 US-09-506-855-43 Sequence 1, Appli
35 31.4 8.9 3182 4 US-08-971-395-1 Sequence 1, Appli
36 31.4 8.9 3183 1 US-08-413-135-1 Sequence 1, Appli
37 31.4 8.9 246240 2 US-08-724-394A-20 Sequence 20, Appli
38 31.4 8.9 246240 2 US-08-724-394A-21 Sequence 21, Appli
39 31.4 8.9 246240 2 US-08-724-394A-22 Sequence 22, Appli
40 31.2 8.9 534 4 US-09-593-995-6 Sequence 6, Appli
41 31 8.8 500 3 US-09-141-000-2 Sequence 2, Appli
42 31 8.8 1697 1 US-08-181-271A-104 Sequence 104, App
43 31 8.8 1697 1 US-08-449-315-104 Sequence 104, App
44 31 8.8 1697 1 US-08-444-803-104 Sequence 104, App
45 31 8.8 1697 1 US-08-449-043-104 Sequence 104, App

ALIGNMENTS

RESULT 1
US-09-334-220-3
; Sequence 3, Application US/09334220
; Patent No. 6323177
; GENERAL INFORMATION:
; APPLICANT: St. Jude's Children's Research Hospital
; APPLICANT: Curran, Thomas
; APPLICANT: D'Arcangelo, Gabriella
; TITLE OF INVENTION: INTERACTION OF REELIN WITH VERY LOW
; TITLE OF INVENTION: DENSITY LIPOPROTEIN (VLDL) RECEPTOR FOR SCREENING AND
; TITLE OF INVENTION: THERAPIES
; FILE REFERENCE: 2427/0F704
; CURRENT APPLICATION NUMBER: US/09/334, 220
; CURRENT FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 11673
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-334-220-3

Query Match 71.1%; Score 249.4; DB 4; Length 11673;
Best Local Similarity 57.7%; Pred. No. 1.7e-72;
Matches 202; Conservative 79; Mismatches 69; Indels 0; Gaps 0;
QY 1 GARCARTGYGCGNACNATNATGCGGNAAYGCGNCTNACNTTTCGTCGAGCCGTACGGCCCT 60
DB 970 GAGCAGTGTGCGACCATCATCATGCGCAATGCTGTCTACCTTCTGTGAGCCGTACGGCCCT 1029
QY 61 MGNGARTNACNACNACNTGYTTNAAACNACNACNACNACNACNACNACNACNACNACNACNACN 120
DB 1030 CGAGAGTGTGCGACCATCATGCGTGTGACACACACACAGCATCTGTCTCCCTCCAGTTTCCATT 1089
QY 121 GGNWSNGNWSNTGYMGNTTYSNTAYWSNGAYCCNWSNATNACNACNACNACNACNACNACN 180
DB 1090 GGGTCAGGATCATCTCGATTTAGTTACTCTGACCCAGCATCATCTGTGTATAGCCCAAG 1149
QY 181 AAYAAACNACNACNACNACNACNACNACNACNACNACNACNACNACNACNACNACNACNACN 240
DB 1150 AACATACCGCTGATGATGATTCAGTTCAGGAGAAATATAGAGCCCTTCAATGTGAGGACA 1209
QY 241 GTNATNACATNTNTAYTNTCCNACNACNACNACNACNACNACNACNACNACNACNACNACN 300
DB 1210 GTATCCACATCTCTACCTCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1269
QY 301 AACARGAYWSNTNMGNTGNGARGCTNTAYGARGCNTGYTGGGNYT 350
DB 1270 AACAGGACAGCCTGCGAGTGGTGTGAGGTGTATGAGGCGCTGCTGGGCGCT 1319

```
RESULT 2
US-09-334-220-4
; Sequence 4, Application US/09334220
; Patent No. 6323177
; GENERAL INFORMATION:
; APPLICANT: St. Jude Children's Research Hospital
; APPLICANT: Curran, Thomas
; APPLICANT: D'Arcangelo, Gabriella
; TITLE OF INVENTION: INTERACTION OF REELIN WITH VERY LOW
; TITLE OF INVENTION: DENSITY LIPOPROTEIN (VLDL) RECEPTOR FOR SCREENING AND
; TITLE OF INVENTION: THERAPIES
; FILE REFERENCE: 2427/0F704
; CURRENT APPLICATION NUMBER: US/09/334,220
; CURRENT FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows version 3.0
; SEQ ID NO 4
; LENGTH: 11580
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-334-220-4

Query Match          66.6%; Score 233.8; DB 4; Length 11580;
Best Local Similarity 56.0%; Pred. No. 2.9e-67;
Matches 196; Conservative 74; Mismatches 80; Indels 0; Gaps 0;

QY 1 GARCARTGYGNCACNATNATGCAAGGNAAYGCNGTNCACNTTYTGYGARCCTTAYGNCN 60
Db 860 GAACAGTGTGGCGGATATGATGCGCAATGCGGTCACTTCTGTGAACCATATGGCCCA 919
QY 61 MGNGARYTNACNACNACNTGYTNAAYACNACNACNCSNGTNTYTCARTTYSNATN 120
Db 920 CGAAGCTGATACCAAGCCCTTAATACAAACACAGCTTCTGCTCCCAATTTCCATT 979
QY 121 GGNWSNGNNSNTGYMNTTYSNTAYWSNGAYCCNWSNATNACNCTNWSNTAYGCAAR 180
Db 980 GGGTCAGTTTCATGTCGCTTAGTATTCAGACCCAGCATCATCGTGTATATGCAAG 1039
QY 181 AAYAAACNCGNCAATGATNCARYTNGARAATNMGNCNCSNNAAYGTNWSNACN 240
Db 1040 AATAACTCTCGGACTGGATGATGAGTACAGTAAAGAAATAGAGCCCTTCCCAATGTCAGACA 1099
QY 241 GTNATNCAYATNTYNTAYTNCNGARGCNAARGNGGARGNCTNCTARTTYCARTGG 300
Db 1100 ATCATCATATCTCTACCTTCTGAGGACGCCAAGGGAGAGATGTCCAATTTCACTGG 1159
QY 301 AARCAGAYWSNTNMGNTNGNGARGTNTAYGARGCNTGYTGGGNYT 350
Db 1160 AAGCAGGAAAATCTTCGTGTAGTGAAGTGTATGAAGCCTGCTGGGCCTT 1209

RESULT 3
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match          11.2%; Score 39.2; DB 1; Length 7218;
Best Local Similarity 7.4%; Pred. No. 0.037;
Matches 25; Conservative 140; Mismatches 172; Indels 0; Gaps 0;

QY 1 GARCARTGYGNCACNATNATGCAAGGNAAYGCNGTNCACNTTYTGYGARCCTTAYGNCN 60
Db 1448 GAAGAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1389
QY 61 MGNGARYTNACNACNACNTGYTNAAYACNACNACNCSNGTNTYTCARTTYSNATN 120
Db 1388 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1329
QY 121 GGNWSNGNNSNTGYMNTTYSNTAYWSNGAYCCNWSNATNACNCTNWSNTAYGCAAR 180
Db 1328 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1269
QY 181 AAYAAACNCGNCAATGATNCARYTNGARAATNMGNCNCSNNAAYGTNWSNACN 240
Db 1268 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1209
QY 241 GTNATNCAYATNTYNTAYTNCNGARGCNAARGNGGARGNCTNCTARTTYCARTGG 300
Db 1208 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1149
QY 301 AARCAGAYWSNTNMGNTNGNGARGTNTAYGARGCNTGYTGGGNYT 337
Db 1148 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1112

RESULT 4
US-09-412-554A-3
; Sequence 3, Application US/09412554A
; Patent No. 6355788
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell
; APPLICANT: Ellsworth, Jeff L.
; TITLE OF INVENTION: FOLLISTATIN RELATED PROTEIN 2FSTA2
; FILE REFERENCE: 98-50
; CURRENT APPLICATION NUMBER: US/09/412,554A
; CURRENT FILING DATE: 1999-10-05
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2949
; TYPE: DNA
```

```

Db 325  AYTYYTCNGRTGRTYYTCYTCCANCRCRTCNAARNSDATRTTTCNCCKRAAANARNSWRA 266
QY 97  GCNWSNGTNTNCARTYYWSNATNGGNWSNGNWSNTGYMGNTTYWSNTAYWSNGAYCCN 156
Db 265  ADATYTGYNARCATYTCRTGNADATNGCNARNGRTTGNCCYTTTGTGTAYTGYTGNG 206
QY 157  WSNATNACNGTNWSNTAYGCNAARAAYAAACNGCNGAYTGGATNCARYTNGARAARATN 216
Db 205  GNSWNARNSWTYTGTGGNGNARNARRAARTYTTTNCRTGTNGGNARRCAVYTGTYGDATNS 146
QY 217  MGNCNCNCSNAAAYGTNWSNACNGTNATNCAYATNTNTAYTNTCCNGARGAGCNAAR 276
Db 145  WNAENGTYTGNARYTTTTRTNARNARYTNTARNNSWYTCYGTGTNTNACYTGNCKYTGTCRA 86
QY 277  GGNGARWSNGTNCARTYTCARTGGAARCARGAYWSNTNMGNTNGGNGAR 327
Db 85  ADATDATNARYTTNARTCNARNSWRAADATNGTNGTNSWNGCNARNARNA 35

RESULT 6
US-09-251-645-11
; Sequence 11, Application US/09251645
; Patent No. 6281413
; GENERAL INFORMATION:
; APPLICANT: Kramer, Vance C.
; APPLICANT: Morgan, Michael K.
; APPLICANT: Anderson, Arne R.
; APPLICANT: Hart, Hope
; APPLICANT: Warren, Gregory W.
; APPLICANT: Dunn, Martha
; APPLICANT: Chen, Jeng S.
; TITLE OF INVENTION: NOVEL INSECTICIDAL TOXINS FROM PHOTORHABDUS LUMINESCENS
; FILE REFERENCE: GCG1963/A
; CURRENT APPLICATION NUMBER: US/09/251,645
; CURRENT FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 37948
; TYPE: DNA
; ORGANISM: Photorhabdus luminescens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (15171)..(18035)
; OTHER INFORMATION: orf5
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (23768)..(31336)
; OTHER INFORMATION: hph2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (31393)..(35838)
; OTHER INFORMATION: orf2
US-09-251-645-11

Query Match 9.7%; Score 34; DB 4; Length 37948;
Best Local Similarity 30.9%; Pred. No. 9.4;
Matches 43; Conservative 26; Mismatches 70; Indels 0; Gaps 0;

QY 76  ACNTGYTYTNAAYACNACNACNGNWSNGTNTNCARTYYWSNATNGGNWSNGNWSNTGY 135
Db 14454  ACCCATGCTCATGATCCCATTTGGTCGAGAAATCAAGGTCACTACCCGCAAAGGCTGGTTC 14513
QY 136  MGNTTYSNTAYWSNGAYCCNWSNATNACNGTNWSNTAYGCNAARAAYAAAYACNGCNGAY 195
Db 14514  CGTCGAACCTTGTTCACTCCCTTGTTTACTGTCTCAATGAAGATGAATGACACAGCTACT 14573
QY 196  TGGATNCARYTNGARAARA 214
Db 14574  GAGGTGAAGGTAAGAAGA 14592

```



```

; CURRENT APPLICATION NUMBER: US/09/128,155
; CURRENT FILING DATE: 1998-08-03
; EARLIER APPLICATION NUMBER: US 60/091,650
; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: US 60/054,646
; EARLIER FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 152331
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(152331)
; OTHER INFORMATION: n = A,T,C or G
US-09-128-155-16

Query Match          9.5%   Score 33.2; DB 3; Length 152331;
Best Local Similarity 27.5%; Pred. No. 62;
Matches 41; Conservative 31; Mismatches 77; Indels 0; Gaps 0;

Qy      114 YWSNATNGGNWNSGNNSTCYMGNTYTWSNAYWSNGAYCCNNSNATNACNGTNWSNTA 173
       ::: ||| :::: |:: |::: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |::
Db      8925 TAGAAAGGGGCACCAGTATTGGCATCTTTCTCACCOCAGGACCACTTGCTGCCATGTCTCTAG 8984

Qy      174 YGCNAARAAYAAYANGCNONGAYTGATNCARYTNGARAAARTNMNGNCNCCNWSNAAVGT 233
       || |:: |||| |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |:: |::
Db      8985 TGCTGCACAGCACCACTGGCCTTGCTGTCGTACTGTTTTGAATAGTCTCGCTGCACAAGAAGGC 9044

Qy      234 NWSNACNGTNATNCAAYATNYTNTAYTNC 262
       : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      9045 ACCAAGGGGCCAGCACCACTGCTCTGCTGTC 9073
```

```

RESULT 11
US-09-134-001C-543
; Sequence 543, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GPC-007
; CURRENT APPLICATION NUMBER: US/09/134.001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 543
; LENGTH: 957
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-543

Query Match          9.48; Score 33; DB 4; Length 957;
Best Local Similarity 25.89; Pred. No. 0.69;
Matches 66; Conservative 40; Mismatches 150; Indels 0; Gaps 0;

Qy 28 AAYGCGTGNACNTTGTGGARCCNTAYGGNCNMGNGARYTNACNACNACNTGYGTNAAY 87
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 60 AACITTGTACATACTACGGGCGGTGATCCCTTCAAAGAATTACAACAAATATTATAAAT 119

Qy 88 ACNACNACNGCNSNGTNYTNCARTYTNWSNATNGNNSNGNWSNTGYMGNTTYWSNTAY 147
    | : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 120 GGCCCTTCTTCAAGTAAATTTAATACCTGCTTTTCAGCGCGCATGGCTTGCAGTAGTAAT 179

Qy 148 WSGAYCCNWSNATNACNGTWNSTAYGCNAAARAAYAAAYACNGCNGAYTGGATNCARYTN 207
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 180 GACAAACACTCCCTTCCTATCTCCATCTCCACAAATACATATTGATGGTACTGGCTCTAC 239

```

Qy	208	GARAARATNMGNCNCNWSNAAAYGTNWSACNGTNAATNTCAATNYNTTAYTYNCCNGAR	267
		: : : :	
Db	240	GCTAATTATGGCGGCGTTGCTGTTTTAAAATAATTATCATCAAGATATTGATCGCAT	299
		: : : :	
Qy	268	GARGCNAARGNGGARW	283
		: : : :	
Db	300	TATGCCCTAGTAAGCAA	315
		: : : :	
RESULT 12			
US-09-397-992A-6/C			
; Sequence 6, Application US/09397992A			
; Patent No. 6329175			
; GENERAL INFORMATION:			
; APPLICANT: Conklin, Darrell			
; APPLICANT: Grant, Francis J.			
; APPLICANT: Rixon, Mark W.			
; APPLICANT: Kindsvogel, Wayne			
; TITLE OF INVENTION: Interferon-epsilon			
; FILE REFERENCE: 98-46			
; CURRENT APPLICATION NUMBER: US/09/397,992A			
; CURRENT FILING DATE: 1999-09-16			
; PRIOR APPLICATION NUMBER: 60/101,012			
; PRIOR FILING DATE: 1998-09-18			
; PRIOR APPLICATION NUMBER: 60/118,578			
; PRIOR FILING DATE: 1999-02-05			
; PRIOR APPLICATION NUMBER: 60/142,766			
; PRIOR FILING DATE: 1999-07-08			
; NUMBER OF SEQ ID NOS: 33			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 6			
; LENGTH: 624			
; TYPE: DNA			
; ORGANISM: Artificial Sequence			
; FEATURE:			
; OTHER INFORMATION: This degenerate sequence encodes the amino acid			
; OTHER INFORMATION: sequence of SEQ ID NO:5.			
; NAME/KEY: variation			
; LOCATION: (1)...(624)			
; OTHER INFORMATION: n is any nucleotide			
US-09-397-992A-6			
Query Match 9.3%; Score 32.8; DB 4; Length 624;			
Best Local Similarity 6.2%; Pred. No. 0.54;			
Matches 18; Conservative 98; Mismatches 175; Indels 0; Gaps 0;			

```

Qy 37 ACNTYTYGYGARCNTYAYGGNCCNMGNARYTNACNACNACNTGYTTYNAAYACNACHACN 96
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 325 AYTTTTCNGRTGTGTTTCTVCCCAACCRCTCNARNNSWDATRTTNGCNCCKKAARNSWRA 2666
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy 97 GCNWSNGTNYTNCARTTYWSNATNGNWSNGNWSNTGYMGNTTYWSNTAYWSNGAYCCN 1566
   :: | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 265 ADATYTGTGNARCATYTCRTGNARDATNGCNRANGCRTGCNCCTTYTGTARTYTYGTGNG 2060
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy 157 WSNATNACNGTNNWSNTAYGCNAAARAAYAACNGCNGAYTGGATNCARTYTNAGARAATN 2160
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 205 GNSWNARNSWTTYTGNGGNARNARRAARTYTTNCKRTGNGGNARRCAYTGYTGDATA 1468
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy 217 MGNCGNCCNWSNAAIYGTNWSNACNGTNTATNCAYATNTNTAYTTCNCGNARGARGCCNAAR 2767
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 145 WNARNGTYTGNARTYTRTTNARNARYTTNARNSWTYCYTGRTTNACYTGNCKYTGTYGRA 86
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Qy 277 GGNCARWSNGTNCARTTYTCARTGGAARCARGAYSNTNMGNGTNGNGGAR 327
   : : : : : : : : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 85 ADATDATNARYTTNARKTTCNARNSWRAADATNGTNGTSNWNGCNARNARNA 35
   : : : : : : : : | : | : | : | : | : | : | : | : | : | : | : | : | :

RESULT 13
US-09-106-194 -11
; Sequence 11, Application US/09106194
; Patent No. 6262234
; GENERAL INFORMATION:
; APPLICANT: Holloway, James
```

```

; APPLICANT: Jelinek, Laura
; APPLICANT: Durnam, Diane
; APPLICANT: Blumberg, Hal
; TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR POLYPEPTIDE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/106,194
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Leith, Debra K
; REGISTRATION NUMBER: 32,619
; REFERENCE/DOCKET NUMBER: 96-11
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6674
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2067 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-106-194-11

```

```

Query Match 9.3%; Score 32.6; DB 4; Length 2067;
Best Local Similarity 9.7%; Pred. No. 1.9;
Matches 29; Conservative 92; Mismatches 177; Indels 1; Gaps 1;

QY 1 GACARTGYGNCACNATNATGCAAGGNAAYGCNGTACNATTTTTCGACCCNTAYGGCCN 60
Db 749 GNGNWSNRCARCAAAATGCGNSNTAYWSNTGYCCNMGNCARMGNAAYTGYTTNATHG 808

QY 61 MGNGARTNACNACNATGYYTNAAYAC-NACNACGNCNWSNGTNYTNCARTTYWSNAT 119
Db 809 AYGNCACNAAYMGNAAYMGNTGYCARCAYTCYMGNTNCARAAFTGYTNGCNYTNGGNA 868

QY 120 NGNWSNGNWSNGTGYMGNTTYWSNTAYWSNGAYCCNWSNATNACNGTNSNTAYGCNAA 179
Db 869 TGWSNMGNGAYGCGNTNAARTTYGGNMGNTGWSNAAARCAARCMGNGAYWSYNTAYG 928

QY 180 RAAYAAACNGCNGAYTGGATNGARAAATNGARAAATNGCNCNWSNAAAYGTTNWSNAC 239
Db 929 CNGARTNCAARCAACATCARGMGTNCARGARCMGNCARCAARWSNGNGARG 988

QY 240 NGTNCATYATNTYNTAYTNGCNGARGAAGGNGGNGNWSNATNACRTTYCAT 298
Db 989 CNGARGCNYTNGCMGNGTNTAYWSNWSNATHWSNAAAYGGNTYWSNAAAYTNAAY 1047

```

```

RESULT 14
US-08-816-693A-1/c
; Sequence 1, Application US/08816693A
; Patent No. 5874241
; GENERAL INFORMATION:
; APPLICANT: Takahashi, Joseph S
; APPLICANT: Turek, Fred W

```

```

; APPLICANT: Pinto, Lawrence H
; TITLE OF INVENTION: Clock Gene and Gene Product
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dressler, Rocky, Milnamow & Katz
; STREET: Two Prudential Plaza, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/816,693A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5874241thrup, Thomas E
; REGISTRATION NUMBER: 33,268
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5400
; TELEFAX: 312-616-5460
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7498 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 389..2954
; US-08-816-693A-1

```

```

Query Match 9.3%; Score 32.6; DB 2; Length 7498;
Best Local Similarity 24.8%; Pred. No. 6.2;
Matches 62; Conservative 43; Mismatches 145; Indels 0; Gaps 0;

QY 52 TAYGGNCCNMNGARYTNACNACNACNTGYTTAAAYACNACNACNWSNGTNYTNCAR 111
Db 1156 TCTGACAGTAGTACAAAACAACTCTATCTTCAATAAGAGGCTATGTGCGTCTAT 1097

QY 112 TTWSNATNGNWSNGNWSNTGYMGNTTYWSNTATWSNGAYCCNWSNATNACNGTWSN 171
Db 1096 AGTTCCTTCAAAACCACTGTGTGTGAAGTTGATACACTGGTTAAAGATTAAATTC 1037

QY 172 TAYGCNAAARAAAYACNCGNGAYTGGATNCARYTNGARAAATNMGNCNWSNAA 231
Db 1036 TATAAATCTACATATTCATAGTGGATGGCTCTTTGGGTCTATTGTTCTCGAAGCAT 977

QY 232 GTNWSNACNATNACAYATNTYNTAYTNCNGARGCARGCNGGNGGNGTNCAR 291
Db 976 GTGACAAACAAATCTAACTGATTTTGTGATTTTAACTACTCAGGGGTTAATGAGTCACT 917

QY 292 TTYCARTGGA 301
Db 916 TTCAGCAGA 907

```

```

RESULT 15
US-08-885-291-1/c
; Sequence 1, Application US/08885291A
; Patent No. 6057125
; GENERAL INFORMATION:
; APPLICANT: Takahashi, Joseph S.
; APPLICANT: Turek, Fred W.
; APPLICANT: Pinto, Lawrence H.
; TITLE OF INVENTION: CLOCK GENE AND GENE PRODUCT
; FILE REFERENCE: 0290-5
; CURRENT APPLICATION NUMBER: US/08/885,291A

```

; CURRENT FILING DATE: 1997-06-30
; EARLIER APPLICATION NUMBER: 08/816,693
; EARLIER FILING DATE: 1997-03-13
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 7498
; TYPE: DNA
; ORGANISM: Mus musculus
US-08-885-291-1

Query Match 9.3%; Score 32.6; DB 3; Length 7498;
Best Local Similarity 24.8%; Pred. No. 6.2;
Matches 62; Conservative 43; Mismatches 145; Indels 0; Gaps 0;
QY 52 TAYGGNCCNMGARYTNACNACNACNTGYTNAAYACNACNACNCGNWSNGTNYTNCAR 111
Db 1156 TCTGACAGTAGCTACAAACAACTCTATCTTCATAAGAGGCCCTATGTGTGCGTTGTAT 1097
QY 112 TTYWSNATNGNWSNGNWSNTGYMNTTYWSNTAYWSNGAYCCNWSNATNACNGTNWSN 171
Db 1096 AGTTCCTTCAAAACCATTTGCTGTGAAGTTGATACACTGGTTAAAGATTTAAATTTCC 1037
QY 172 TAYGCNAAARAAYACNCGNAYVTGGATNCARYTNGARARATNMGNCNCCNWSNAA 231
Db 1036 TATAAATCTCACATATTTCATAGGTGGATGGCTCTTTGGGTCTATTGTTCTCGAAGCAT 977
QY 232 GTNWSNACNGTNATNCAYATNTYATYTNCCNCGARGCNAAARGGNGARWSNGTNCAR 291
Db 976 GTGACAAACAAGATTTCTAACTGATTTTGTGATTTTAACTACTCAGGGGTTAATGAGTCACT 917
QY 292 TTYCARTGGA 301
Db 916 TTCAGCAGA 907

Search completed: November 7, 2002, 02:19:44
Job time : 156 secs

THIS PAGE BLANK (USPTO)

Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	249.4	71.1	351	10	US-09-697-430B-1	Sequence 1, Appli	
2	233.8	66.6	11580	10	US-09-880-107-3436	Sequence 3436, Ap	
3	90.2	25.7	499	10	US-09-864-761-14497	Sequence 14497, A	
4	78.2	22.3	214	10	US-09-864-761-31037	Sequence 31037, A	
5	53.4	15.2	413	10	US-09-864-761-11808	Sequence 11808, A	
6	44.8	12.8	87	10	US-09-864-761-28379	Sequence 28379, A	
7	35.4	10.1	485	10	US-09-864-761-10999	Sequence 10999, A	
8	35.2	10.0	696	10	US-09-825-561A-8	Sequence 8, Appli	
9	34.8	9.9	1905	10	US-09-922-488-3	Sequence 3, Appli	
10	34.6	9.9	593	10	US-09-864-761-7539	Sequence 7539, Ap	
11	34.6	9.9	594	10	US-09-917-800A-1003	Sequence 1003, Ap	
12	34.6	9.9	4068	10	US-09-962-436-295	Sequence 295, App	
13	34.4	9.8	12932	10	US-09-764-847-1132	Sequence 1132, Ap	
14	33.6	9.6	921	10	US-09-925-300-223	Sequence 223, App	
15	33.4	9.5	1278	10	US-09-815-242-6280	Sequence 6280, Ap	
16	33.4	9.5	180216	10	US-09-835-232-6	Sequence 6, Appli	
17	33.2	9.5	560	10	US-09-864-761-32024	Sequence 32024, A	
18	33.2	9.5	1989	10	US-09-864-761-15510	Sequence 15510, A	
19	33.2	9.5	2379	12	US-10-006-867-123	Sequence 123, App	

Query Match 71.1%; Score 249.4; DB 10; Length 351;
Best Local Similarity 57.7%; Pred. No. 1.3e-65;
Matches 202; Conservative 79; Mismatches 69; Indels 0; Gaps 0;

QY	1	GARCATGTGGNACNATNATGCAYGAGNAYGCGNGTNACNTTYTGYGARCNNTAYGNCN	60
Db	1	GAGCAGTGTGGCCACCATCATCATGCAATGCTGTCACTTCTGTGAGCGGTACGCCCTT	60
QY	61	MGNARYTNACNACNACNTGYTNAAYACNACNACNGCNWSNCTNTNCARTYWSNATN	120
Db	61	CGAGAGCTGACCAACCACATGCCTCAACACACAAACAGCATCTGCTCCTCACATTTTCCATT	120
QY	121	GGNWSGNGNWSNTGYMGNTTYSNTAYWSNGAYCCNWSNATNACNGTNNWSNTAYGCNAAR	180
Db	121	GGCTGAGGATCATGTCGATTTAGTTACTCTGAGCCCAAGCATCACTGTGTCATACAGCCAAG	180
QY	181	AAYAAACCCNCAYTGGATNCARYNGARAAARATNMGNGCCNCCNWSNAAAYGTNNSNACN	240
Db	181	AACCAATACCCCTGATTGGATTACGCTGGAGAAAATTAGAGCCCTTTCCAAATGTGAGCA	240
QY	241	GTNATNCAYATNTNTAYTYNCCNGARGAGCCNAAAGGNGARWSNCTNCARTTYTCARTGG	300

Db 241 GTCATCCACATCTCTGTACCTCCCGAGGAAGCCAAAGGGGAGAGCTGCAGTTCAGTGG 300

Qy 301 AARCAAGAYNSNTNMGNTGNGARGNTATYARGCNTGYTGGGCNTY 350

Db 301 AACAGACAGCCTGCCAGTGGTGGTGGAGTGTATGAGCCCTGTGGGCCCT 350

RESULT 2

```

US-09-880-107-3436
; Sequence 3436, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherif, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3436
; LENGTH: 11580
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U97916
US-09-880-107-3436

```

Query Match 66.6%; Score 233.8; DB 10; Length 11580;
Best Local Similarity 56.0%; Pred. No. 1.2e-59;
Matches 196; Conservative 74; Mismatches 80; Indels 0; Gaps 0;

[illegible][illegible]

RESULT 3
US-09-864-761-14497/c
; Sequence 14497, Application US/09864761
; Patent NO. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBE
 TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
 FILE REFERENCE: Aeomica-X-1
 CURRENT APPLICATION NUMBER: US/09/864,761
 CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: US 09/632,366
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: GB 24263.6
 PRIOR FILING DATE: 2000-10-04
 PRIOR APPLICATION NUMBER: US 60/236,359
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: PCT/US01/00666
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00667
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00664
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00669
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00665
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00668
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00670
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/774,203
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
 SEQ ID NO 14497
 LENGTH: 499
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AC000121.1
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.62
 US-09-864-761-14497

Query Match	25.7%;	Score 90.2;	DB 10;	Length 499;
Best Local Similarity	57.5%;	Pred. No. 1.5e-17;		
Matches 77: Conservative	27;	Mismatches 30;	Indels 0;	Gaps 0;

Qy	217	MGNGCNCNWSNAAYGTINWSNACNGTNATNCAYATNTTAYTNTCCNGARGARGCNAAR	376
		: : : : : : : : : : :	
Db	487	AGAGCCCTTCCAATGTCACCAACATCCATCTCTACCTCTCTGAGGACGCCAA	428
Qy	277	GGNGRNSGNTNCARTTYCARTGGAARCARGAYWSNYTNNMGNTNGNGARGGTNTAYGAR	336
		: : : : : : : : : : :	
Db	427	GGGAGAAATCTCCAATTTCAGTGGAGCAGGAAATCTTCGTGTAGTGAAGTGTATGAA	368
Qy	337	GCNTGYTGGGCNYT	350
		: : : : : : : : : : :	
Db	367	GCCTGCTGGGCCTT	354

RESULT 4
US-09-864-761-31037/c
; Sequence 31037, Application US/09864761
; Patent No. US20020048763A1

Best Local Similarity 31.4%; Pred. No. 0.47;
Matches 54; Conservative 27; Mismatches 91; Indels 0; Gaps 0;

Qy 160 ATNACNGTNSNTAYGCNNAARAAAYACNCGNCAYTGGATNCARYTNGARAARATWGN 219
Db 303 ACAAAAATTAAGACAAAGTTACTGACTTCAGTTTCCTTTGATTGATTAAGAAACA 244
Qy 220 GCNCCNNSAAYGTNSNACNGTATNCAYATNTAYTNCNCGARGCNAARGN 279
Db 243 ACACATTGCGTTGAACAATTTATTACACATCTTGATGTCTTAAGGACGACGAAGA 184
Qy 280 GARNSGTNCARTTYCARTGGAARCAARGAYWSNTNMGNTGNGGARGTNT 331
Db 183 GGGAAATTGTTATTTCCACTGGAAGAAGAGTGTGTGTCATGCCAGGGCT 132

RESULT 8

US-09-825-561A-8
; Sequence 8, Application US/09825561A
; Patent No. US20020137677A1
; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy A.
; APPLICANT: No. US20020137677A1ak, Julia E.
; APPLICANT: West, James W.
; APPLICANT: Presnell, Scott R.
; APPLICANT: Holly, Richard D.
; APPLICANT: Nelson, Andrew J.
; TITLE OF INVENTION: SOLUBLE ZALPHAL1 CYTOKINE RECEPTORS
; FILE REFERENCE: 00-22
; CURRENT APPLICATION NUMBER: US/09/825,561A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/194,731
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/222,121
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 696
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Degenerate polynucleotide sequence of IL-2Rgamma
; OTHER INFORMATION: polypeptide as shown in SEQ ID NO:4
; NAME/KEY: misc_feature
; LOCATION: (1)...(696)
; OTHER INFORMATION: n = A,T,C or G

Query Match 10.0%; Score 35.2; DB 10; Length 696;
Best Local Similarity 12.1%; Pred. No. 0.72;
Matches 28; Conservative 68; Mismatches 136; Indels 0; Gaps 0;

Qy 114 YWSNATNGNWSNGNWSNTGYMGNTTYSNTAYWSNGAYCCNWSNATNACNGTWSNNTA 173
Db 418 YTNCAAYRTNWSNGARWSNCARTYNGARYTNAAYTGGAAAYAMGNTTYTNAAYCAY 477
Qy 174 YGCNNAARAAAYACNCGNCAYTGGATNCARYTNGARAARATNMGNCNCCNWSNAYGT 233
Db 478 TGYTNGARCAAYTNGTNCARTAYTNGNACNGAYTGGAYCAYSWNTGACNGARCAWSN 537
Qy 234 NWSNACGTNATNCAYATNTAYTNCNCGARGCNAARGNGNGTNCARTT 293
Db 538 GTNCAYTAYMGNCAAYARTTYSNTNCCNWSNGTNGAYGNCARARMGNTAYACNTTY 597
Qy 294 YCARTGGAARCAARGAYWSNTNMGNTGNGGARGTNTAYGARGCNTGYTGG 345
Db 598 MGNCTNMGNSMGNMTTYAAYCCNYTNTGYGNSNCGNCARCAYTGGWSNG 649

RESULT 9

US-09-922-488-3
; Sequence 3, Application US/09922488

; Patent No. US20020119553A1
; GENERAL INFORMATION:
; APPLICANT: Thayer, Edward C.
; APPLICANT: Webster, Phillipa J.
; TITLE OF INVENTION: Human Secreted Protein, Zzpl
; FILE REFERENCE: 00-54
; CURRENT APPLICATION NUMBER: US/09/922,488
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/222,814
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/260,512
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1905
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: degenerate nucleotide sequence
; NAME/KEY: misc_feature
; LOCATION: 6, 9, 12, 15, 18, 21, 24, 30, 36, 39, 42, 45, 48, 51, 54,
; LOCATION: 57, 60, 63, 66, 69, 72, 75, 78, 84, 90, 96, 99, 102, 105,
; LOCATION: 111, 123, 132, 141, 144, 147, 153, 156, 159, 162, 168, 171,
; LOCATION: 174, 183, 186, 198, 204, 213, 225, 243, 246, 249, 252
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: 255, 264, 267, 270, 276, 279, 288, 291, 300, 303, 315, 318,
; LOCATION: 327, 330, 333, 345, 348, 351, 354, 360, 363, 366, 372, 375,
; LOCATION: 384, 387, 390, 399, 405, 411, 414, 417, 420, 423, 429, 435,
; LOCATION: 438, 441, 444, 447, 456, 459, 462, 465, 468, 474, 477
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: 480, 486, 489, 492, 495, 498, 504, 507, 513, 516, 519, 525,
; LOCATION: 531, 534, 537, 540, 546, 549, 555, 558, 561, 567, 570, 573,
; LOCATION: 576, 579, 582, 585, 588, 591, 594, 597, 600, 603, 606, 609,
; LOCATION: 612, 615, 618, 624, 633, 636, 639, 654, 663, 675, 678
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: 684, 687, 705, 708, 711, 714, 720, 723, 732, 735, 738, 741,
; LOCATION: 744, 753, 765, 768, 786, 789, 795, 798, 810, 816, 819, 822,
; LOCATION: 825, 837, 843, 852, 855, 858, 861, 864, 876, 879, 882, 888,
; LOCATION: 894, 897, 900, 912, 915, 921, 924, 927, 930, 936, 939
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: 942, 951, 957, 963, 966, 978, 981, 984, 993, 996, 999, 1008,
; LOCATION: 1011, 1014, 1023, 1041, 1044, 1047, 1050, 1068, 1071, 1077,
; LOCATION: 1080, 1086, 1089, 1095, 1098, 1107, 1113, 1116, 1122, 1131,
; LOCATION: 1134
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: 1143, 1146, 1155, 1158, 1167, 1170, 1173, 1176, 1179, 1182,
; LOCATION: 1185, 1191, 1197, 1200, 1203, 1206, 1209, 1212, 1218, 1221,
; LOCATION: 1227, 1239, 1245, 1248, 1257, 1272, 1278, 1281, 1284, 1287,
; LOCATION: 1290, 1296, 1299, 1305, 1311, 1314, 1317, 1320, 1326
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: 1329, 1335, 1341, 1344, 1347, 1350, 1353, 1368, 1371, 1374,
; LOCATION: 1377, 1380, 1386, 1398, 1407, 1413, 1416, 1422, 1428, 1437,
; LOCATION: 1443, 1449, 1452, 1461, 1464, 1467, 1473, 1476, 1479, 1482,
; LOCATION: 1491, 1503, 1509, 1512, 1515, 1518, 1524, 1527, 1530
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: 1536, 1539, 1542, 1548, 1551, 1554, 1557, 1560, 1563, 1566,
; LOCATION: 1572, 1581, 1584, 1587, 1590, 1599, 1602, 1605, 1608, 1614,
; LOCATION: 1620, 1623, 1626, 1632, 1635, 1638, 1641, 1644, 1647, 1653,
; LOCATION: 1656, 1659, 1662, 1665, 1671, 1680, 1683, 1686, 1689
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: 1701, 1704, 1707, 1710, 1713, 1716, 1719, 1722, 1734, 1740,
; LOCATION: 1749, 1752, 1755, 1758, 1761, 1764, 1770, 1776, 1782, 1785,
; LOCATION: 1788, 1791, 1794, 1797, 1800, 1806, 1809, 1812, 1815, 1818,
; LOCATION: 1821, 1824, 1827, 1830, 1833, 1836, 1839, 1842, 1848

```
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: 1851, 1857, 1860, 1863, 1866, 1872, 1878, 1887, 1896, 1902
; OTHER INFORMATION: n = A,T,C or G
US-09-922-488-3

Query Match          9.9%; Score 34.8; DB 10; Length 1905;
Best Local Similarity 12.1%; Pred. No. 2.2;
Matches 34; Conservative 68; Mismatches 180; Indels 0; Gaps 0;

QY 5 ARTGYGACNATNATGAYCGNAAACGNGTNCACNTTGTGAGCCNTAYGGCCNMGNG 64
Db 1275 HGTNMGNTNTNMGNGARCCGNTCAAGTNGRTNGNTNTNCAACGNGNAYCC 1334
QY 65 ARYTACNACNACNTGYTTAAACNACNACGNGCNGWNGNTNTNCACTTWSNATNGNW 124
Db 1335 NAAATNTNTNTNTNCAACATGTTGGGNGCNCNWSNCCNAAAYCCNTTTCARCA 1394
QY 125 SNGGNWSNTGYMGNTTWSNTAYWSNGAYCCNWSNATNACNGTNTWSNTAYGCNNAARAAYA 184
Db 1395 RCCNCARTGCCNATHYTNWSNGAYGNGTGYCCNTTAAAGGNGAYWSNTAYMGNACNCA 1454
QY 185 AYACNGCNGAYTGATNCARTNCARARATNMGNGCNCNWSNAAAYGTNWSNACNGTNA 244
Db 1455 RATGTTGTCNTGAYGGGNGCNCNCCNTTTCARWSNCAATAYCARMGNTTYACNGTNGC 1514
QY 245 TNCAYATNTNTAYTTNCCNARGCARGCNAARGGNGARWSNG 286
Db 1515 NACNTTTCNTNTNGAYWSNGNWSNCAARMGNGCNYTNG 1556

RESULT 10
US-09-864-761-7539
; Sequence 7539, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 7539
; LENGTH: 593
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC018583.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 7.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.9
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.8
US-09-864-761-7539

Query Match          9.9%; Score 34.6; DB 10; Length 593;
Best Local Similarity 29.3%; Pred. No. 0.96;
Matches 58; Conservative 29; Mismatches 111; Indels 0; Gaps 0;

QY 41 TTTGYGARCCTAYGGCCNMGNGARYTNACNACNACNTGYTTNAAAYACNACNACNGCW 100
Db 89 TCATTTAGACTTCTAGTCTTATTACCTAAGAAATATATTTCTTAAGACTATAGCTGCCA 148
QY 101 SNGTNTNCARTTYWSNATNGNWSNCGNWSNTGYMNTTYSNTAYWSNGAYCCNWSNA 160
Db 149 TAGATAGTGATTTCTCTCAAGAAATCTGGCAAAAGTAAGTTGAAACCTTCTCTGGGGAAGAT 208
QY 161 TNACNGTNSNTAYGCNNAARAAYACNCGNGAYTGATNCARTYTNGARAARATNMGNG 220
Db 209 TCACCATTCAGTAGACAGTAAGAAAATTTGTGATTCATTGGAAGAGGCTCAAGATAACTA 268
QY 221 CNCCNWSNAAAYGTNWSNA 238
Db 269 CACTAACAGAGTTTGA 286

RESULT 11
US-09-917-800A-1003
; Sequence 1003, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
```


OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (885)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (895)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (911)
OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-223

Query Match 9.6%; Score 33.6; DB 10; Length 921;
Best Local Similarity 28.3%; Pred. No. 2.8;
Matches 60; Conservative 34; Mismatches 118; Indels 0; Gaps 0;
Qy 139 TTYNSNTAYNSNGAYCCNWSNATNACNGTNSNTAYCNCNARAAYAYACNGCNGAYTGG 198
Db 652 TTCAGTGATTTAGATCAGAGAACGATTGGAAGTGATCCCAAGGTAGAGCAACAGCTGCT 711
Qy 199 ATNCARYTNGARAARATNMGNCNCCNWSNAAAYGTNNSNACNGTNAATNAYATNTAY 258
Db 712 AACACAAACCTCAGCTAGTAGAAACCGAAGGCCCTCAACTTTTGCCTATGCAGATT 771
Qy 259 YTNCCNGARGCARGCNAAGGNGRWSNGTNCARTTYCARTGGAARCARCARGAYWSNTNMGN 318
Db 772 AATACTAACAGGAGCAAGAGTGCAATTTTACAAGTCCCCCAACAGAGGAAACGGTTGGG 831
Qy 319 GTNGNGARGTNTAYGARGCNTGYTGGGCNTY 350
Db 832 TTCAGCACAGTGTAAAGGNTGTTTTGCCTT 863

RESULT 15

US-09-815-242-6280
Sequence 6280, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011a
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 6280
LENGTH: 1278
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: CDS

LOCATION: (1)...(1278)
US-09-815-242-6280

Query Match 9.5%; Score 33.4; DB 10; Length 1278;
Best Local Similarity 25.7%; Pred. No. 4.2;
Matches 58; Conservative 38; Mismatches 130; Indels 0; Gaps 0;
Qy 31* GCNGTNACNTTYTGYGARCCNTAYGGNCCNMGNGARVTNACNACNACNTGYTTNAAVACN 90
Db 707 GCCGTACCGCTCGTGAAGCGCGCTGGGACACAGTTTCAGTAATTTTGGCCTGAATACG 766
Qy 91 ACNACNCCNWSNGTNTNCAARTTYWSNATNGNWSNNGNWSNNTGYMGNNTTYWSNAYWSN 150
Db 767 ATCCCTCCGATCGTGTCTATCTTCTCTACCTGTGGCGTGTCTGTCTGTCTGTCTGTCTAAGCC 826
Qy 151 GAYCCNWSNATNACNGTNSNTAYCNCNARAAYAYACNCGNGAYTGGATNCARYTNGAR 210
Db 827 TGTTTGTCAATTAACCGCTGCTGCGGATGCGGTGGGCGTGGGCGCTTAAAGCCGCTGCGTG 886
Qy 211 AARATNMGNCNCCNWSNAAAYGTNWSNACNGTNAATNCAATNTNT 256
Db 887 AAGATGAAATCGCCTGCGGTTCGCTGGGCTTAAAGCCGCTGCTAT 932

Search completed: November 7, 2002, 03:29:52
Job time : 94 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2002, 00:46:05 ; Search time 2216 Seconds
(without alignments)
2565.260 Million cell updates/sec

Title: US09897438BK-2
Perfect score: 351
Sequence: 1 GARCARTGYGGNACNATNAT.....AYGARGCNTGYTGGGNCYTN 351

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1:	em_estba:*
2:	em_esthum:*
3:	em_estin:*
4:	em_estnu:*
5:	em_estov:*
6:	em_estpl:*
7:	em_estro:*
8:	em_htc:*
9:	gb_est1:*
10:	gb_est2:*
11:	gb_htc:*
12:	gb_est3:*
13:	gb_est4:*
14:	gb_est5:*
15:	em_estfun:*
16:	em_estom:*
17:	gb_gss:*
18:	em_gss_hum:*
19:	em_gss_inv:*
20:	em_gss_pln:*
21:	em_gss_vrt:*
22:	em_gss_fun:*
23:	em_gss_mam:*
24:	em_gss_mus:*
25:	em_gss_other:*
26:	em_gss_pro:*
27:	em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	179	51.0	315	12	BE938667
c 2	96	27.4	600	12	BG803882
c 3	66.4	18.9	635	10	BB248113
c 4	55	15.7	568	12	BF387628
c 5	46.6	13.3	803	17	CNS018FX
c 6	44.4	12.6	935	17	CNS002X1
					AL098239 Drosophil

7	42.8	12.2	707	17	CNS04ALQ	AL282023 Tetraodon
8	42.8	12.2	1101	17	CNS00KK2	AL077673 Drosophil
c 9	42.6	12.1	855	17	CNS011PE	AL100556 Drosophil
10	42.2	12.0	1101	17	CNS0106X	AL098595 Drosophil
c 11	41.6	11.9	894	17	BH705629	BH705629 BOHWF68TR
12	41.6	11.9	1101	17	CNS00GDR	AL070260 Drosophil
13	41.4	11.8	978	17	CNS02JBA	AL199999 Tetraodon
c 14	41	11.7	1101	17	CNS00KK2	AL077673 Drosophil
c 15	40.4	11.5	746	12	BF694109	BF694109 602082720
c 16	40.4	11.5	844	13	BM400822	BM400822 5009-0-79
17	40.4	11.5	1101	17	CNS0037Q	AL064465 Drosophil
c 18	40.4	11.5	1101	17	CNS014T4	AL104578 Drosophil
c 19	40.2	11.5	803	17	CNS018FX	AL109287 Drosophil
20	40.2	11.5	1101	17	CNS00FOO	AL070854 Drosophil
21	40	11.4	972	17	CNS02BJJ	AL189928 Tetraodon
c 22	40	11.4	1101	17	CNS0142D	AL104803 Drosophil
23	39.8	11.3	555	17	AZ937550	AZ937550 2M0195H13
24	39.8	11.3	556	17	AZ937529	AZ937529 2M0195D15
c 25	39.8	11.3	732	17	CNS03NYB	AL252668 Tetraodon
c 26	39.8	11.3	1101	17	CNS00CNT	AL059413 Drosophil
27	39.6	11.3	894	12	BG441805	BG441805 GA_Ea001
28	39.4	11.2	650	12	BG404162	BG404162 602420202
c 29	39.4	11.2	1101	17	CNS0106D	AL098575 Drosophil
30	39.2	11.2	980	17	CNS00JMO	AL076572 Tetraodon
c 31	39.2	11.2	987	17	CNS035IN	AL071398 Drosophil
32	39.2	11.2	1100	17	CNS00G3S	AL228776 Tetraodon
c 33	39	11.1	306	17	AZ579300	AZ579300 1M0363N07
34	39	11.1	787	17	CNS010B7	AL098749 Drosophil
c 35	39	11.1	793	17	CNS012YX	AL102195 Drosophil
36	39	11.1	1101	17	CNS00L4E	AL067752 Drosophil
37	39	11.1	1101	17	CNS00KT0	AL077995 Drosophil
c 38	38.8	11.1	975	17	CNS00L8Q	AL068009 Drosophil
39	38.8	11.1	1033	17	CNS013FB	AL102785 Drosophil
c 40	38.8	11.1	1101	17	CNS00B2I	AL058538 Drosophil
c 41	38.8	11.1	1101	17	CNS00FOO	AL070854 Drosophil
42	38.6	11.0	624	14	BQ392162	BQ392162 NISC_mq22
c 43	38.6	11.0	787	17	CNS010B7	AL098749 Drosophil
44	38.6	11.0	1039	17	CNS006H7	AL064351 Drosophil
c 45	38.6	11.0	1204	17	CNS016E2	AL106628 Drosophil

ALIGNMENTS

RESULT 1
BE938667/c
LOCUS
DEFINITION QV0-TN0084-180800-342-a08 TN0084 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE938667
VERSION BE938667.1 GI:10466224
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 315)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICK Human Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tlj=et2-QV0-TN0084-180
 800-342-a08&t3=2000-08-18&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 315.

FEATURES

Location/Qualifiers
 source

1. .315
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="TN0084"
 /dev_stage="Adult"
 /note="Organ: testis_normal; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."
 BASE COUNT 84 a 64 c 78 g 89 t
 ORIGIN

Query Match 51.08; Score 179; DB 12; Length 315;
 Best Local Similarity 53.88; Pred. No. 5.6e-41;
 Matches 149; Conservative 61; Mismatches 67; Indels 0; Gaps 0;
 QY 74 CNACTGYTTNAAACACACACAGCTTCTGTCCTCCAAATTCATTTGGGTGAGGTTCAT 133
 Db 312 CTAGGCCCTTATACAAACACAGCTTCTGTCCTCCAAATTCATTTGGGTGAGGTTCAT 253
 QY 134 GYMGNTTYSNTAYNSGAYCCNWSNATNACNGTNSNTAYGNCNAAAYAAACNGCNG 193
 Db 252 GTCCGCTTAGTTATTCAGACCCAGCATCATCGTGTGTATATGCGCAAGTAACCTCTCGG 193
 QY 194 AYTGGATNCARYTNGARAARTNMGNCNWSNAAAYGTTNWSNACNGTNCAYATNY 253
 Db 192 ACTGGATTACAGTAGAGAAAATTAGAGCCCTTCCCAATGTGAGCAATCATCCATATCC 133
 QY 254 TNTAYTNCNGARGAGCNAARGNGGARGNSNTNCARTTYCARTTGGGAARCAARGAYWSNY 313
 Db 132 TCTACCTTCCTGAGAGCCCAAGGGGAGAGATGTCCAAATTCAGTGGAGCAGGAAGATC 73
 QY 314 TNMGNTNGNGARGTNTAYGARGCNTGTGGGCNYT 350
 Db 72 TTCGTGTAGGTGAAGTGTATGAAGCCTGCTGGGCCTT 36

RESULT 2
 BG803882/c
 LOCUS
 DEFINITION 600 bp mRNA linear EST 20-DEC-2001
 0243-03 Mouse E14.5 retina lambda ZAP II Library Mus musculus cDNA,
 mRNA sequence.

ACCESSION BG803882
 VERSION BG803882.1 GI:17950794
 KEYWORDS EST.
 SOURCE house mouse.

ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 600)

REFERENCE
 AUTHORS Mu, X., Zhao, S., Perashad, R., Hsieh, T.-F., Scarpa, A., Wang, S.W.,
 White, R.A., Beremand, P.D., Thomas, T.L., Gan, L. and Klein, W.H.
 Gene expression in the developing mouse retina by EST sequencing
 and microarray analysis

TITLE Nucleic Acids Res. 29 (24), 4983-4993 (2001)
 JOURNAL 21671825
 MEDLINE

COMMENT Contact: Klein WH
 Department of Biochemistry and Molecular Biology
 University of Texas M.D. Anderson Cancer Center
 Box 117, 1515 Holcombe Blvd., Houston, TX 77030, USA
 Tel: 713 792 3546
 Fax: 713 790 0329.

FEATURES

Location/Qualifiers
 source

1. .600
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone_lib="Mouse E14.5 retina lambda ZAP II Library"
 /tissue_type="neural retina"
 /dev_stage="embryonic day 14.5 post-fertilization"
 /note="Vector: pAMP10 (Gibco); Cloned unidirectionally.
 Primer: Oligo dT. RNA Isolation: cytoplasmic RNA preps
 (Manniat); Cloning Technique: CUA Cloning (CloneAmp,
 Life Technologies); Average insert size: 1.8 Kb;
 Insertion site: TAGTCCACTGAATCTGAGT-->. Other
 information regarding entire library may be found at
 http://pga.swmed.edu/Data/Libraries/microarray_cdna_librar
 ies.htm."

BASE COUNT 127 a 167 c 152 g 152 t 2 others
 ORIGIN

Query Match 27.48; Score 96; DB 12; Length 600;
 Best Local Similarity 58.68; Pred. No. 1.1e-16;
 Matches 78; Conservative 30; Mismatches 25; Indels 0; Gaps 0;

QY 218 GNGCNCNWSNAAAYGTNWSNACNGTNCAYATNYTNTAYTNCNGARGARCCNAARG 277
 Db 476 GAGCCCTTCCAATGTGAGCAGCATCATCCATCTGTTACCTCCCGAGGAGCCAAAG 417
 QY 278 GNGARWSNGTNCARTTYCARTTGGGAARCAARGAYWSNYTNGTNGGARGCTNTAYGARG 337
 Db 416 GGGAGCGGTGCAGTTCAGTGGAAACAGACAGCCCTCGGAGTGGGTGATGAGG 357
 QY 338 CNTGYTGGGCNYT 350
 Db 356 CCGTGTGGGCCT 344

RESULT 3

BB248113
 LOCUS
 DEFINITION 635 bp mRNA linear EST 23-OCT-2001
 musculus cDNA clone A730023J04 3', mRNA sequence.

ACCESSION BB248113
 VERSION BB248113.2 GI:16355611
 KEYWORDS EST.
 SOURCE house mouse.

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

On Jul 6, 2000 this sequence version replaced gi:8940859.
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,

VERSION AL109287.1 GI:5629591
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 803)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : ww.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelOBAC11.
FEATURES Location/Qualifiers
source 1..803
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACN13P04"
/clone_lib="DrosBAC"
/plasmid="pBelOBAC11"
/note="end : T7"
BASE COUNT 119 a 113 c 128 g 138 t 305 others
ORIGIN
Query Match 13.3%; Score 46.6; DB 17; Length 803;
Best Local Similarity 12.2%; Pred. No. 0.03;
Matches 30; Conservative 102; Mismatches 114; Indels 0; Gaps 0;
Qy 92 CNACGNCNWSGNTYNTNCARTYWSNATNGNWSNGNWSNTGYMNTYWSNTAYWSNG 151
Db 756 CCBSSCNKSKBKCKBKCKBKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCKCK 697
Qy 152 AYCCNWSNATNACNTNWSNTAYGCNAAAYAAACNGCNGAYTGGATNCARYTNGARA 211
Db 696 KCCCCCBKWKKKKKKKKAKTCCWAMCAAAAMMAAMWYAATTAACABBKAWAMADA 637
Qy 212 ARATNMGNCNCCNSNAAYGTNWSNACNGTNCATNCAYATNTYNTAYTNCNCGARGARG 271
Db 636 KKKMHKKCKCKMTADAKKKCKMAKKDKADACCMKAAKKAAKCKCKKKKAKKKA 577
Qy 272 CNAARGNGARWSNCTNCARTYTCARTGGAARCARGAYWSNTYNNMGNTNGNGARGTNT 331
Db 576 AKTKAKKKKKNNNNKMAAKAWAKKAKKACDCAADAAADAAVKKDAKAAKWKAKSGRAK 517
Qy 332 AYGARG 337
Db 516 KKGKKK 511
RESULT 6
CNS002X1 935 bp DNA linear GSS 26-JUL-1999
LOCUS Drosophila melanogaster genome survey sequence SP6 end of BAC
DEFINITION BACN02A14 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL098239
VERSION AL098239.1 GI:5609850
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 935)
AUTHORS Genoscope.

TITLE Direct Submission
JOURNAL Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelOBAC11.
FEATURES Location/Qualifiers
source 1..935
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACN02A14"
/clone_lib="DrosBAC"
/plasmid="pBelOBAC11"
/note="end : SP6"
BASE COUNT 95 a 36 c 129 g 307 t 368 others
ORIGIN
Query Match 12.6%; Score 44.4; DB 17; Length 935;
Best Local Similarity 11.9%; Pred. No. 0.15;
Matches 38; Conservative 123; Mismatches 158; Indels 0; Gaps 0;
Qy 20 TCCAYGGNAAYCCGNTNACNTTYTGYGARCCTNATYGGNCCNNGARAYTNACNACNACNT 79
Db 617 TTMNMGTTMSVTAHSGMMMTGCTTMMKMMKSGMTTTCVTGTMGTATTMTAMMMMM 676
Qy 80 GYTTNAAAYACNACNACNCGNWSNGTNTNCARTTYWSNATNGNWSNGNWSNTGYMNT 139
Db 677 STGGCMMTMTACMNMNMGTRMMHMMHMMHMMHMMHMMHMMHMMHMMHMMHMMHMMH 736
Qy 140 TYWSNTAYWSNGAYCCNWSNATNACNGTNSNTAYGCNAAAYAAAYACNGCNGAYTGA 199
Db 737 TRACTVAMTMTVMTBATGGTGCAGCTGMMMTACGMGTVMMTTGMGTGAVGAGTKTT 796
Qy 200 TNCARYTNGARAARATNMGNCNCCNWSNAAAYGTNWSNACNCTNATNCAYATNTYNTAY 259
Db 797 KTTATTTTGTGAMGARRMTVGVDTTGHAVMTTMMRRRSGKTTGGARRTNGGAATAMT 856
Qy 260 TNCNCGARGACNAAARGNGARWSNGTNCARTTYTCARTGGAARCARGAYWSNTYNNMG 319
Db 857 KRMTHHTTKTTGTTTHSACAGVWSGMMVHBTGNGMTTGHVAGTGVMMMMHMMHMMHMM 916
Qy 320 TNGNGARGTNTAYGARGC 338
Db 917 KGRMTVMACMMSAMGMRM 935
RESULT 7
CNS04ALQ 707 bp DNA linear GSS 21-MAY-2000
LOCUS Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
DEFINITION 095B14 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL282023
VERSION AL282023.1 GI:8020357
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 707)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
TITLE Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence

[illegible]

RESULT 15	BF694109/c	746 bp	mRNA	linear	EST 23-DEC-2000
LOCUS	602082720F1 NIH_MGC_81	Homo sapiens	cdna	clone	IMAGE:4247069 5',
DEFINITION	BF694109	BF694109 human. human.			
ACCESSION	BF694109	BF694109.1 GI:11979517			
VERSION	EST.	human.			
KEYWORDS	EST.	human.			
SOURCE	human.	human.			
ORGANISM	Homo sapiens	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
TITLE	NIH-MGC http://mgc.nci.nih.gov/.	NIH-MGC http://mgc.nci.nih.gov/.			
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)	National Institutes of Health, Mammalian Gene Collection (MGC)			
COMMENT	Unpublished (1999)	Unpublished (1999)			
	Contact: Robert Strausberg, Ph.D.	Contact: Robert Strausberg, Ph.D.			
	Email: cgabps-femail.nih.gov	Email: cgabps-femail.nih.gov			
	Tissue Procurement: CLONTECH Laboratories, Inc.	Tissue Procurement: CLONTECH Laboratories, Inc.			
	CDNA Library Preparation: CLONTECH Laboratories, Inc.	CDNA Library Preparation: CLONTECH Laboratories, Inc.			
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
	DNA sequencing by: Incyte Genomics, Inc.	DNA sequencing by: Incyte Genomics, Inc.			
	Clone distribution: MGC clone distribution information can be	Clone distribution: MGC clone distribution information can be			
	found through the I.M.A.G.E. Consortium/LLNL at:	found through the I.M.A.G.E. Consortium/LLNL at:			
	http://image.llnl.gov	http://image.llnl.gov			
	Plate: LLC1063 row: g column: 06	Plate: LLC1063 row: g column: 06			
	High quality sequence stop: 643.	High quality sequence stop: 643.			

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4247069"
/clone_lib="NIH_MGC_81"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: muscle (skeletal); Vector: pDNR-LIB
(Clontech); Site_1: Sfil (ggcgctgcggc); Site_2: Sfil
(ggcattatggc); 5' and 3' adaptors were used in cloning
as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGC-3',
and 3' adaptor sequence:
5'-ATTCTAGCGCGCGCGGCAGCATG-CT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.55 kb (range 1.0-4.0 kb). 15/15 colonies contained
inserts by PCR. This library was enriched for full-length
clones and was constructed by Clontech Laboratories (Palo
Alto, CA)."
```

[illegible]

Db 343 TCAATGGCCACATTAAATAACCAAAATAATTAACACACAGAAAGTTAAACCTACAATACATT 284
Qy 167 TNWNTAYGCNAARAAYACNGCNGAYTGGATNCARYTNGARAARATNMGNGCNCNW 226
Db 283 TCACACATCTACTGTGTAAACCACACAAATTTTAAAAAATTTTAAAAATATTTTATACCT 224
Qy 227 SNAAYGTNWSNACNGTNAATNCAYATNTAYTNCNCNGARGARGCNAARGGNGARWSNG 286
Db 223 TAAATATTAAAAAATATTTTAAAAATGGCTATGCAAGAAAAAAGGAAAAAGAAATACAT 164
Qy 287 TNCARTTYCARTGGAARCARGAYWSNTNMGNGTNGNGAR 327
Db 163 TTCATTTACCATATATAACACACACTTTAAAAATGGGTAGAA 123

Search completed: November 7, 2002, 02:17:01
Job time : 2223 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 6, 2002, 22:05:00 ; Search time 2212 Seconds
(without alignments)
856.633 Million cell updates/sec

Title: US-09-897-438B-2

Perfect score: 117

Sequence: 1 EQCGTTHMGNAVTFCEPYGP.....FQWKQDSLRGVEYACWAL 117

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Word size: 1

Total number of hits satisfying chosen parameters: 32308016

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-Q=/cgn2_1/USPTO.spool/US09897438/runat_06112002_101956_11811/app_query.fasta_1.263
-DB=EST -QFMT=fastap -SUFFIX=oligo.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09897438_@CGN1_1_1716_@runat_06112002_101956_11811 -NCPU=6 -ICPU=3
-NO_XLPAY -NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
c 1	44	37.6	600	12	BG803882
c 2	31	26.5	635	10	BG803882 0243-03 M
c 3	27	23.1	315	12	BB248113
c 4	13	11.1	568	12	BE938667
c 5	9	7.7	717	17	BF387628 UI-R-CAL-
c 6	8	6.8	260	17	AQ688885 nxb0078D
c 7	8	6.8	308	10	AZ010548
c 8	8	6.8	312	12	AW114381 rs51810.Y
c 9	8	6.8	354	10	BF642099 NF039F071
c 10	8	6.8	366	17	AW911732
c 11	8	6.8	371	14	AZ791276 2M0041F09
c 12	8	6.8	381	17	C47012 C47012 YuJ1
c 13	8	6.8	381	17	AZ033696
c 14	8	6.8	399	17	AQ087955
c 15	8	6.8	400	10	AQ030475
c 16	8	6.8	400	10	AW696758
c 17	8	6.8	411	10	AW789434
c 18	8	6.8	411	10	AV803611
c 19	8	6.8	442	10	AW655178
c 20	8	6.8	451	9	AA758580
c 21	8	6.8	456	12	BB825159
c 22	8	6.8	457	13	BF775130
c 23	8	6.8	475	17	BM105816
c 24	8	6.8	476	9	AQ994137
c 25	8	6.8	478	12	AL369916
c 26	8	6.8	479	10	BF639389
c 27	8	6.8	489	17	AW668785
c 28	8	6.8	491	10	AQ603779
c 29	8	6.8	493	17	AW684418
c 30	8	6.8	501	9	AQ345388
c 31	8	6.8	503	17	AL367517
c 32	8	6.8	509	9	AQ669357
c 33	8	6.8	519	13	AU207398
c 34	8	6.8	524	12	BI589981
c 35	8	6.8	528	12	BF641085
c 36	8	6.8	534	17	BE808142
c 37	8	6.8	539	17	AQ562548
c 38	8	6.8	546	12	AQ773077
c 39	8	6.8	549	10	BE754467
c 40	8	6.8	549	13	AW669102
c 41	8	6.8	550	17	BI775248
c 42	8	6.8	551	17	AL765609
c 43	8	6.8	552	9	AQ298583
c 44	8	6.8	557	12	AI997627
c 45	8	6.8	566	12	BE756494
					BF005427

ALIGNMENTS

RESULT 1
BG803882/c
LOCUS 0243-03 Mouse E14.5 retina lambda ZAP II Library Mus musculus cdNA, EST 20-DEC-2001
DEFINITION mRNA sequence.
ACCESSION BG803882
VERSION BG803882.1 GI:17950794
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 600)
REFERENCE Mu, X., Zhao, S., Pershad, R., Hsieh, T.-F., Scarpa, A., Wang, S.W., White, R.A., Beremand, P.D., Thomas, T.L., Gan, L. and Klein, W.H.

TITLE	Gene expression in the developing mouse retina by EST sequencing
JOURNAL	Nucleic Acids Res. 29 (24), 4983-4993 (2001)
MEDLINE	21671825
COMMENT	Contact: Klein WH Department of Biochemistry and Molecular Biology University of Texas M.D. Anderson Cancer Center Box 117, 1515 Holcombe Blvd., Houston, TX 77030, USA Tel: 713 792 3646 Fax: 713 790 0329.

FEATURES	Location/Qualifiers	source
	l. .600	
	/organism="Mus musculus"	
	/db_xref="taxon:10090"	
	/clone_lib="Mouse E14.5 retina lambda ZAP II Library"	
	/tissue_type="neural retina"	
	/dev_stage="embryonic day 14.5 post-fertilization"	
	/note="Vector: pAMP10 (Gibco); Cloned unidirectionally. Primer: Oligo dT. RNA Isolation: cytoplasmic RNA preps (Manniatls); Cloning Technique: CUA Cloning (CloneAmp, Life Technologies); Average insert size: 1.8 Kb; Insertion site: TACGCCACGAACTCGAGTC--->. Other information regarding entire library may be found at http://pga.swmed.edu/Data/Libraries/microarray/cdna_library.htm "	

BASE COUNT	127 a	167 c	152 g	152 t	2 others
ORIGIN					
Alignment Scores:					
Pred. No.:	6.58e-37		Length:		600
Score:	44.00		Matches:		44
Percent Similarity:	100.00%		Conservative:		0
Best Local Similarity:	100.00%		Mismatches:		0
Query Match:	37.61%		Indels:		0
DB:	12		Gaps:		0

US-09-897-438B-2 (1-117) x BG803882 (1-600)

74	AlapProSerAsnValSerThrValIleHisIleLeuTyrLeuProGluGluAlaTysGly	93
QY		
474	GCCTTCCTCAATGTGAGCAGAGTCATCCACATCTCTACCTCCCGAGAGGCCAAAGGG	415
Db		
94	GluSerValGlnphedInTrpLysGlnAspSerLeuArgValGlyGluValTyrGluAla	113
QY		
414	GAGAGCGTCGACTTCAGTGGAAACAGGACAGCCCTCGGAGTGGGTGAGGTGATGAGGCC	355
Db		
114	CysTrpAlaLeu	117
QY		
354	TGCTGGGCCCTG	343
Db		

RESULT 2	LOCUS	635 bp	mRNA	linear	EST 23-OCT-2001
BB248113	BB248113	RIKEN full-length enriched, 7 days neonate cerebellum Mus musculus	cdna clone A730023J04	3'	mRNA sequence.

ACCESSION BB248113
VERSION BB248113.2 GI:16355611
KEYWORDS EST.

SOURCE	house mouse.
ORGANISM	Mus musculus
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 635)
 Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,
 Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda
 M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M.,
 Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki
 D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,
 Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toyota, T.,
 Muramatsu, M. and Hayashizaki, Y.
 RIKEN Mouse ESTs (Arakawa, T., et al. 2001)

RIKEN Mouse ESTs (Arakawa, T., et al. 2001)

JOURNAL
Unpublished (2001)
On Jul 6, 2000 this sequence version replaced qi:8940859.
COMMENT

Contact: Yoshihide Hayashiraki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@gsc.riken.go.jp,
 URL: <http://genome.gsc.riken.go.jp/>
 Carninci, P., Shibata, Y., Hayatsu, M. and Hayashizaki, Y.,
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Sugahara, Y., Shibata, K., Itoh
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. *Genome Res.* 10 (10), 1617-1630 (2000)
 wagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
 Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipillar sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)

Kondo, S., Shingawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y.

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cdDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

FEATURES	SOURCE
1. <i>General</i>	
2. <i>Specific</i>	
3. <i>Other</i>	

```

1. 0033
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="A730023J04"
/clone_lib="RIKEN full-length enriched, 7 days neonate
cerebellum"
/tissue_type="cerebellum"
/dev_stage="7 days neonate"
/lab_host="DH10B"
/notes="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGATCCAGAGCTCTTTTTTTTTTTTNN 3'], cDNA was
prepared by using trihalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 20.0 and subtraction to Rot = 459.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGAGATTCGAGTTAATAAATATCCCCCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pluescript KS(+) after bulk excision from Lambda
phage.

```

	FLC 1.		
	182 a	122 c	118 g
BASE COUNT			213 t
ORIGIN			

Alignment Scores:		
Pred. No.:	6.55e-23	Length: 635
Score:	31.00	Matches: 31
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0

Query Match: 26.50% Indels: 0
DB: 10 Gaps: 0

US-09-897-438b-2 (1-117) x BB248113 (1-635)

QY 11 AlaValThrPheCysGluProTyrGlyProArgGluLeuThrThrCysLeuAsnThr 30
|||||

Db 3 GCTGTCCACCTTCTGTGAGCGGTACGCGCTCGAGAGCTGACCACCATGCGCTGAACACA 62
|||||

QY 31 ThrThrAlaSerValLeuGlnPheSerIleGly 41
|||||

Db 63 ACACAGCATGTCTCCAGTTTCCCATGGT 95
|||||

RESULT 3
LOCUS BE938667/c 315 bp mRNA linear EST 02-OCT-2000

DEFINITION QV0-TN0084-180800-342-a08 TN0084 Homo sapiens cDNA, mRNA sequence.

ACCESSION BE938667

VERSION BE938667.1 GI:10466224

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bai,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=st2-QV0-TN0084-180
800-342-a08&t3=2000-08-18&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 315.
Location/Qualifiers
1. 315
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="TN0084"
/dev_stage="Adult"
/note="Organ: testis normal; Vector: puc18; Site.1: SmaI;
Site.2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 84 a 64 c 78 g 89 t

ORIGIN

Alignment Scores:
Pred. No.: 5,21e-19 Length: 315
Score: 27.00 Matches: 27
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 23.08% Indels: 0
DB: 12 Gaps: 0

1000

US-09-897-438B-2 (1-117) x BE938667 (1-315)

QY 28 LeuAsnThrThrAlaSerValLeuGlnPheSerIleGlySerGlySerCysArgPhe 47
|||||

Db 304 CTTAATACACAAACAGCTTCTGCTCCCAATTTTCCATTGGGTGAGGTCATGCTCGCTTT 245
|||||

QY 48 SerTyrSerAspProSerIle 54
|||||

Db 244 AGTTATTACAGACCCAGCATC 224
|||||

RESULT 4
LOCUS BF387628/c 568 bp mRNA linear EST 27-NOV-2000

DEFINITION UI-R-CAL-bbs-e-06-0-UI.s1 UI-R-CAL Rattus norvegicus cDNA clone

ACCESSION BF387628

VERSION BF387628.1 GI:11372452

KEYWORDS EST.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene
discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NCI site
and the oligo-dT track served to identify it as a clone from the
normalized pons library cDNA library preparation: M.B. Soares Lab
Clone distribution: clones will be available through Research
Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA=yes.

FEATURES Location/Qualifiers
1. 568
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-CAL-bbs-e-06-0-UI"
/clone_lib="UI-R-CAL"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker: Site.1: Not I; Site.2: Eco RI; The UI-R-CAL
library is a subtracted library derived from the following
tissues: thalamus, cerebellum, hypothalamus, medulla, pons
, midbrain, cerebral cortex, corpus striatum, testis, and
hippocampus. For a detailed description of the library
from which this clone was derived, please visit our web
site at ratest.eng.uiowa.edu. The subtraction has been
previously described in (Bonaldo, Lennon and Soares,
Genome Research 6:791-806, 1996)
TAG_LIB="UI-R-CAL"
TAG_TISSUE="pons"
TAG_SEQ="AGCACG"

BASE COUNT 116 a 156 c 131 g 165 t

ORIGIN

Alignment Scores:
Pred. No.: 0.00123 Length: 568
Score: 13.00 Matches: 13

us-09-897-438b-2.oligo.rst

us-09-897-438b-2 (1-117) x BE938667 (1-315)

Page 3

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 11.11% Indels: 0
 DB: 12 Gaps: 0

US-09-897-438B-2 (1-117) x BF387628 (1-568)

QY 105 LeuArgValGlyGluValTyrGluAlaCysTrrpAlaLeu 117
 Db 526 CTCGCTGGGTGAGGTGTACGAGCCTGCTGGCCCTG 488

RESULT 5

AQ688885

LOCUS

DEFINITION AQ688885 nbxb0078D03f CUGI Rice BAC Library Oryza sativa genomic clone

260 bp DNA linear GSS 01-JUL-1999

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 717)
 A BAC End Sequencing Framework to Sequence the Rice Genome
 Unpublished (1998)
 Contact: Wing RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 7288
 Fax: 864 656 4293
 Email: twing@clemson.edu
 Seq primer: TAATACGACTCACTATAGG
 Class: BAC ends
 High quality sequence stop: 305.

FEATURES

SOURCE

1..717

/organism="Oryza sativa"

/strain="Japonica"

/cultivar="Nipponbare"

/db_xref="taxon:4530"

/clone="nbxb0078D03f"

/tissue_type="Leaf"

/lab_host="E. coli DH10B"

/notes="Vector: pBeloBAC11; Site_1: HindIII; Site_2: HindIII; Rice is one of two most popular grains in the world. Half of the world population especially those inhabiting highly populated areas of the humid tropics and subtropics, rely on rice as their primary source of carbohydrate. Monocotyledonous rice is a diploid plant (2n=24) with a haploid genome equivalent of 431 Mbp (Arumuganathan and Earle, 1991). The relatively small genome of rice, three times larger than that of Arabidopsis, makes it suitable for genomic studies. In order to facilitate positional cloning, physical mapping and genome sequencing of rice, we have constructed a BAC library from Oryza sativa, Nipponbare variety. The library contains 36,864 clones with an average insert size of 128.5 kb providing 10.9 haploid genome equivalents. The deep coverage allows the isolation a particular sequence with a probability of 99.9 %. Two high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening."

BASE COUNT

ORIGIN

Alignment Scores:
 Pred. No.: 33.1 Length: 717
 Score: 9.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 7.69% Indels: 0
 DB: 17 Gaps: 0

US-09-897-438B-2 (1-117) x AQ688885 (1-717)

QY 41 GlySerGlySerCysArgPheSerTyr 49
 Db 359 GGAAGCGGTTCAATGCCGTTTCTCTAC 385

RESULT 6

AZ010548

LOCUS

DEFINITION AZ010548 RPCI-23-325B10.TJ RPCI-23 Mus musculus genomic clone

260 bp DNA linear GSS 25-FEB-2000

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 260)
 Mus musculus house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akınret, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P., and Fraser, C. M.
 Mouse BAC End Sequences from Library RPCI-23
 Unpublished (1999)
 Other_GSSs: RPCI-23-325B10.TV
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm) or from Resea ch Genetics (info@resgen.com). BAC end page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
 Plate: 325 row: B column: 10
 Seq primer: SP6
 Class: BAC ends.

FEATURES

source

1..260

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="RPCI-23-325B10"

/sex="Female"

/lab_host="DH10B"

/note="Organ: Kidney/brain; Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methyase. Size selected DNA was cloned into the pBACE3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT

ORIGIN

Alignment Scores:
 Pred. No.: 105 Length: 260
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 6.84% Indels: 0
 DB: 17 Gaps: 0

US-09-897-438B-2 (1-117) x AZ010548 (1-260)

QY 41 GlySerGlySerCysArgPheSer 48
 Db 217 GGAAGTGGCTCCTGTAGATTACG 240

RESULT 7
 LOCUS AW114381 308 bp mRNA linear EST 10-MAY-2001
 DEFINITION rs51b10.y1 Sommer Pristionchus Pristionchus pacificus cDNA 5', mRNA sequence.

ACCESSION AW114381
 VERSION AW114381.1 GI:6080719
 KEYWORDS EST.
 SOURCE Pristionchus pacificus.
 ORGANISM Pristionchus pacificus.

REFERENCE 1 (bases 1 to 308)
 AUTHORS Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; Neodiplogasteridae; Pristionchus.

McCarte, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J., Wylie, T., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C., Tsagarelis, W., Ronko, I., Kennedy, S., Maguire, L., Beck, C., Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T., Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

TITLE The Washington Univ. Nematode EST Project, 1999
 JOURNAL Unpublished (1999)
 COMMENT Contact: McCarter JP

The Washington Univ. Nematode EST Project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

The library was constructed by Dr. Ralf Sommer DNA Sequencing by:
 Washington University Genome Sequencing Center
 Contact Dr. Ralf Sommer (ralf.sommer@uebingen.mpg.de) for
 information about this clone.

Putative full length read
 The vector to vector length is 374
 Seq primer: -40RP from Gibco
 High quality sequence stop: 307.

FEATURES
 Location/Qualifiers

1..308
 /organism="Pristionchus pacificus"
 /strain="PS 312"
 /db_xref="taxon:54126"
 /clone_lib="Sommer Pristionchus"
 /sex="predominantly hermaphroditic"
 /dev_stage="mixed stages (embryo to adult)"
 /lab_host="not applicable (host cell line)"
 /note="Vector: Uni-ZAP XR vector (Stratagene); Site_1: 5' EcoRI; Site_2: 3' XhoI; 1st strand cDNA was primed with a XhoI - oligo(dT) primer. Double-stranded cDNA was ligated to EcoRI adaptors digested with XhoI and cloned into XhoI and EcoRI sites. Primary complexity of the library was 10 in the 7th. The library went through one round of amplification."

BASE COUNT 69 a 92 c 38 g 108 t 1 others

ALIGNMENT SCORES
 Pred. No.: 131 Length: 308
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 6.84% Indels: 0
 DB: 10 Gaps: 0

US-09-897-438B-2 (1-117) x AW114381 (1-308)

QY 50 SerAspProSerIleThrValSer 57

Db 272 TCGACCCATCCATTACAGTATCC 295

RESULT 8
 LOCUS BF642099/c

DEFINITION BF642099 312 bp mRNA linear EST 19-DEC-2000
 NF039F07IN1062 Insect herbivory Medicago truncatula cDNA clone sequence.

ACCESSION BF642099
 VERSION BF642099.1 GI:11906257
 KEYWORDS EST.
 SOURCE barrel medic.
 ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

REFERENCE 1 (bases 1 to 312)
 AUTHORS Korth, K., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.

TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation
 JOURNAL Medicago truncatula insect herbivory library
 COMMENT Unpublished (2000)

CONTACT: Korth K
 DEPT. of Plant Pathology
 University of Arkansas

217 Plant Science Building, Fayetteville, AR 72701, USA
 Tel: 501 575 5191

Email: korth@comp.uark.edu
 Fax: 501 575 7601

Insert Length: 312 Std Error: 0.00

Plate: 039 row: F column: 07

Seq primer: TCACACAGGAAACAGCTATGAC.

FEATURES
 Location/Qualifiers

1..312
 /organism="Medicago truncatula"
 /db_xref="taxon:3880"
 /clone_lib="NF039F07IN"
 /clone_lib="Insect herbivory"
 /tissue_type="local and systemic leaves"
 /note="Vector: Lambda zap; Library was produced from fully expanded M. truncatula leaves of plants fed upon by Spodoptera exigua (beet armyworm) for 24 hours. Systemic (undamaged leaves from injured plants) and wounded leaves were harvested and pooled."

BASE COUNT 69 a 69 c 84 g 90 t

ORIGIN

ALIGNMENT SCORES
 Pred. No.: 133 Length: 312
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 6.84% Indels: 0
 DB: 12 Gaps: 0

US-09-897-438B-2 (1-117) x BF642099 (1-312)

QY 74 AlaProSerAsnValSerThrVal 81

LOCUS AW911732

DEFINITION AW911732 354 bp mRNA linear EST 25-MAY-2000
 uf38f01.v1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone sequence.

ACCESSION AW911732
 VERSION AW911732.1 GI:8077314
 KEYWORDS EST.

SOURCE house mouse.
 ORGANISM Mus musculus

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 354)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Other_ESTs: uf38f01.x1
Contact: Robert Strausberg, Ph.D.
Email: ccgaps-r@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:940509
Seq primer: -40RP from Gibco
High quality sequence stop: 332.
FEATURES
source
1..354
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:1513657"
/clone_lib="Soares_mammary_gland_NMLMG"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/notes="vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT73 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Ronaldo."
BASE COUNT 68 a 99 c 118 g 69 t
ORIGIN

Alignment Scores:
Pred. No.: 157 Length: 354
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.84% Indels: 0
DB: 10 Gaps: 0

US-09-897-438B-2 (1-117) x AW911732 (1-354)

QY 39 SerIleGlySerGlySerCysArg 46
Db 7 TCCATCGGCTCTGCAGCTGCAGG 30

RESULT 10
AZ791276/c 366 bp DNA linear GSS 16-FEB-2001
LOCUS
DEFINITION
2M0041F09F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC2M0041F09 F, DNA sequence.
ACCESSION
AZ791276
VERSION
AZ791276.1 GI:12933997
KEYWORDS
GSS.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 366)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.,
and Wright,D. Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

```

```

84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0041 row: F column: 09
Seq primer: CGTTGTAAACGACGCGCAGT
Class: plasmid ends
High quality sequence stop: 366.
FEATURES
source
1..366
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0041F09"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWB42 (gil14732114(gb)AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
BASE COUNT 122 a 72 c 54 g 118 t
ORIGIN

Alignment Scores:
Pred. No.: 164 Length: 366
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.84% Indels: 0
DB: 17 Gaps: 0

US-09-897-438B-2 (1-117) x AZ791276 (1-366)

QY 30 ThrThrThrAlaSerValLeuGln 37
Db 281 ACTACTACACCAAGTGTCTGCAG 258

RESULT 11
C47012/c 371 bp mRNA linear EST 18-OCT-1999
LOCUS
DEFINITION
C47012 Yuji Kohara unpublished cDNA:Strain N2 hermaphrodite embryo
Caenorhabditis elegans cDNA clone yk428h6 5', mRNA sequence.
ACCESSION
C47012
VERSION
C47012.1 GI:2383265
KEYWORDS
EST.
SOURCE
Caenorhabditis elegans.
ORGANISM
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae
; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE
1 (bases 1 to 371)
AUTHORS
Kohara,Y., Mochizuki,T., Tabara,H., Watanabe,H., Sugimoto,A., Sano
,M., Miyata,A. and Nishigaki,A.
Expression map of the C.elegans genome
Unpublished (1996)
CONTACT: Yuji Kohara
Genome Biology Lab.

```


National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
Location/Qualifiers

FEATURES

source

1..371
/organism="Caenorhabditis elegans"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk428h6"
/clone_lib="Yuji Kohara unpublished cDNA:Strain N2
hermaphrodite embryo"
/sex="hermaphrodite"
/dev_stage="embryo"
1..80 g 123 t

BASE COUNT 97 a 71 c 80 g 123 t
ORIGIN

Alignment Scores:
Pred. No.: 166 Length: 371
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.84% Indels: 0
DB: 14 Gaps: 0

US-09-897-438B-2 (1-117) x C47012 (1-371)

QY 43 GlySerCysArgPheSerTyrSer 50

LOCUS
Db 169 GGAAGTTGCGATTTTCATCACTCT 146

RESULT 12

AZ033696/c

LOCUS
DEFINITION
AZ033696 381 bp DNA linear GSS 01-MAR-2000
RNA sequence.
AZ033696
Mus musculus genomic clone RPCI-23-248G7,
DNA sequence.

ACCESSION
VERSION
AZ033696.1 GI:7118183

KEYWORDS
SOURCE
house mouse.

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 381)
Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret
B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
and Fraser,C.M.
Mouse BAC End Sequences from Library RPCI-23
Unpublished (1999)

Other_GSSs: RPCI-23-248G7-TV

TITLE

CONTACT: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 248 row: G column: 7
Seq primer: SP6
Class: BAC ends.

FEATURES

source

1..381
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-248G7"

/clone_lib="RPCI-23"
/sex="Female"
/lab_host="DH10B"
/note="Organ: Kidney/brain; Vector: pBACE3.6; Site_1:
ECORI; Site_2: EORI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EORI and EORI Methylase. Size
selected DNA was cloned into the pBACE3.6 vector at the
ECORI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 107 a 73 c 76 g 125 t
ORIGIN

Alignment Scores:
Pred. No.: 172 Length: 381
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.84% Indels: 0
DB: 17 Gaps: 0

US-09-897-438B-2 (1-117) x AZ033696 (1-381)

QY 30 ThrThrAlaSerValLeuGln 37

Db 112 ACTACTACAGCAAGTGTCTGCAG 89

RESULT 13

AQ087955/c

LOCUS

DEFINITION
HS_2191_B1_D02_MR CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate-2191 Col=3 Row=H, DNA sequence.

ACCESSION
VERSION
AQ087955.1 GI:3456866

KEYWORDS
SOURCE
human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 381)

Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.

Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome

Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

99380589

MEDLINE

COMMENT

Contact: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Sequence Tagged Connector

Plate: 2191 row: H column: 3

Class: BAC ends

High quality sequence stop: 381.

FEATURES

source

1..381
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate-2191 Col=3 Row=H"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
E-Coli DH10B."

BASE COUNT 69 a 81 c 131 g 100 t
ORIGIN

Alignment Scores:

Pred. No.: 172 Length: 381
Score: 8.00 Matches: 8

Percent Similarity: 100.00%
 Best Local Similarity: 100.00%
 Query Match: 6.84%
 DB: 17

Conservative: 0
 Mismatches: 0
 Indels: 0
 Gaps: 0

US-09-897-438B-2 (1-117) x AQ087955 (1-381)

Qy 46 ArgPheSerTyrSerAspProSer 53
 Db 186 AGATTCTCCTACTCTGACCCCTCT 163

RESULT 14
 AQ030475

LOCUS
 DEFINITION HS_2183_A1_G06_MF CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate-2183 Col-11 Row=M, DNA sequence.

ACCESSION AQ030475

VERSION AQ030475.1 GI:3275401

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 399)

AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
 Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
 Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and

scanning the human genome

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

MEDLINE 99380589

COMMENT Contact: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Sequence Tagged Connector

Plate: 2183 row: M column: 11

Class: BAC ends

High quality sequence stop: 399.

Location/Qualifiers

1..399

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="plate-2183 Col-11 Row=M"

/notes="Organ: sperm; Vector: pBelOBAC11; BAC Clones in

E-Coli DH10B"

BASE COUNT 114 a 100 c 57 g 123 t 5 others

ORIGIN

Alignment Scores:

Pred. No.: 183 Length: 399

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 6.84% Indels: 0

DB: 17 Gaps: 0

US-09-897-438B-2 (1-117) x AQ030475 (1-399)

Qy 76 SerAsnValSerThrValIleHis 83

Db 245 TCAAATGTTTCACTGTCTATCCAT 268

RESULT 15

AQ030475

LOCUS

DEFINITION NF108E09St 5', mRNA sequence.

AW696758 400 bp mRNA linear EST 20-DEC-2000

LOCUS

DEFINITION NF108E09St 5', mRNA sequence.

ACCESSION AW696758

VERSION AW696758.2 GI:11934078

KEYWORDS EST.

SOURCE barrel medic.

ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;

Medicago.

REFERENCE 1 (bases 1 to 400)

AUTHORS He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell

,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon

,R.A.

TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation

Medicago truncatula stem library

JOURNAL Unpublished (2000)

COMMENT On Apr 14, 2000 this sequence version replaced gi:7571520.

Contact: Dixon RA

Plant Biology Division

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 221 7302

Fax: 580 221 7380

Email: radixon@noble.org

Insert Length: 738 Std Error: 0.00

Plate: 108 row: E column: 09

Seq primer: TCACACAGGAACAGCTATGAC.

Location/Qualifiers

1..400

/organism="Medicago truncatula"

/db_xref="taxon:3880"

/clone="NF108E09St"

/clone_lib="Developing stem"

/tissue_type="stem"

/dev_stage="Pooled developmental"

/notes="Vector: Lambda Zap; Contains a mixture of

internodal stem segments"

BASE COUNT 93 a 87 c 105 g 115 t

ORIGIN

Alignment Scores:

Pred. No.: 184 Length: 400

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 6.84% Indels: 0

DB: 10 Gaps: 0

US-09-897-438B-2 (1-117) x AW696758 (1-400)

Qy 74 AlaproSerAsnValSerThrVal 81

Db 312 GCTCCTCCCAATGTGTCAACGGTC 289

Search completed: November 6, 2002, 23:35:29

Job time : 2218 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 6, 2002, 18:13:22 ; Search time 2770 Seconds
(without alignments)
1229.252 Million cell updates/sec

Title: US-09-897-438B-2
Perfect score: 625
Sequence: 1 EQCGTIMHGNAVTFCEPYGP.....FQWKQDSLVRGVEVYACWAL 117

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USFTO.spool/US09897438/runat_06112002_101924_11289/app_query.fasta_1.263
-DB=GenEmbl -Qfmt=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09897438.ecgn.1.1.2659 @runat_06112002_101924_11289 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_ov.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_or.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

em_vi.*
29: em_htg_hum.*
30: em_htg_inv.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
1	625	100.0	11673	6	AX305377 Sequence
2	625	100.0	11673	10	U24703 Mus musculus
3	596	95.4	11187	10	AB049473 Rattus no
4	572	91.5	11580	6	AX410790 Sequence
5	572	91.5	11580	9	U79716 Human reell
6	572	91.5	11580	11	G30936 SWSS2926 Er
7	572	91.5	11580	11	G30938 SWSS3176 Er
8	418	66.9	10634	5	AF090441 Gallus ga
9	241	38.6	202764	2	AC128022 Rattus no
10	240	38.4	163985	10	AC121878 Mus muscu
11	240	38.4	163985	2	AC023062 Mus muscu
12	237	37.9	183641	2	AC124933 Rattus no
13	231	37.0	126130	2	AC095877 Rattus no
14	224	35.8	93163	9	HSAC000121 Human BAC
15	144	23.0	185996	2	AC041023 Homo sapi
16	118	18.9	158	4	AF232904 Bos tauru
17	92	14.7	4960	10	MUSBALBC1
18	89.5	14.3	1259	5	AF090951 Lacerata v
19	87.5	14.0	2092	5	AF090843 Emys orbi
20	86.5	13.8	231756	2	AC109258 Mus muscu
21	85.5	13.7	142100	2	AC094067 Rattus no
22	84	13.4	166764	2	AC129106 Homo sapi
23	84	13.4	170606	9	AC018898 Homo sapi
24	84	13.4	171862	2	AC009442 Homo sapi
25	84	13.4	343300	2	AC107296 Homo sapi
26	83	13.3	1845	6	I07862 Sequence 2
27	83	13.3	1845	6	I08324 Sequence 3
28	83	13.3	1845	6	I08327 Sequence 6
29	83	13.3	3408	6	I08074 Sequence 1
30	83	13.3	3408	6	I08323 Sequence 1
31	83	13.3	10519	1	AE001040 Archaeogl
32	81	13.0	2514	8	ANGLUA
33	81	13.0	2514	8	S73370 glucoamylas
34	81	13.0	2602	6	AR198404 Sequence
35	81	13.0	2602	6	AX073733 Sequence
36	81	13.0	2602	8	ANGA01
37	81	13.0	3411	8	ASNGIGII
38	80.5	12.9	110000	2	LMFLCHR12_3
39	80.5	12.9	110000	2	LMFLCHR16_03
40	80	12.8	3411	6	E00315 DNA coding
41	80	12.8	4819	5	MGU49431 Meleagris g
42	80	12.8	19823	2	AC019868 Drosophil
43	80	12.8	43525	3	U88311 Caenorhabd
44	80	12.8	163378	2	AC008336 Drosophil
45	80	12.8	174681	2	AC107156 Rattus no

ALIGNMENTS

RESULT 1

AX305377
LOCUS AX305377 11673 bp DNA linear PAT 11-DEC-2001
DEFINITION Sequence 128 from Patent WO0188188.
ACCESSION AX305377
VERSION AX305377.1 GI:17644926
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Ishikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y.
TITLE Method for examining ischemic conditions
JOURNAL Patent: WO 0188188-A 128 22-NOV-2001;
School Juridical Person Nihon University (JP)
FEATURES
source
1. 11673
/organism="Mus musculus"
/db_xref="taxon:10090"
BASE COUNT 2831 a 2985 c 2985 g 2872 t
ORIGIN
Alignment Scores:
Pred. No.: 6,42e-66 Length: 11673
Score: 625.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-897-438B-2 (1-117) x AX305377 (1-11673)
Qy 1 GluInCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20
|||||
Db 970 GAGCAGTGGCACCACCATGTCATGCAATGCTGTCACCTCTGTGAGCGCTAGCGCCCT 1029
Qy 21 ArgGluLeuThrThrThrCysLeuAsnThrThrThrAlaSerValLeuGlnPheSerIle 40
|||||
Db 1030 CGAGAGCTGACCACACACATGCTGSAACACAAACAGACATGCTGCTCCAGTTTCCATT 1089
Qy 41 GlySerGlySerCysArgPheSerTyrSerAspProSerIleThrValSerTyrAlaLys 60
|||||
Db 1090 GGTGAGATCATGTCGATTTAGTTACTCTGACCCACATCATCTGTCATACGCCAAG 1149
Qy 61 AsnAsnThrAlaAspTrpIleGlnLeuGluLysIleArgAlaProSerAsnValSerThr 80
|||||
Db 1150 AACAAATACCGTGTGATTGATTCAGTGGAGAAATAGAGCCCTTCCAAATGTGAGCACA 1209
Qy 81 ValIleHisIleLeuTyrLeuProGluGluAlaLysGlyGluSerValGlnPheGlnTrp 100
|||||
Db 1210 GTATCCACATCCCTGATCTACCTCCCGAGGAGCAAGGAGGAGCGTCCAGTTCAGTGG 1269
Qy 101 LysGlnAspSerLeuArgValGlyGluValTyrGluAlaCysTrpAlaLeu 117
|||||
Db 1270 AACAGGACAGCGCTGGAGTGGGTGAGGTGTATGAGGCTGCTGGGCCCTG 1320
RESULT 2
MMU24703
LOCUS MMU24703 11673 bp mRNA linear ROD 19-DEC-1997
DEFINITION Mus musculus reelin mRNA, complete cds.
ACCESSION U24703
VERSION U24703.1 GI:2702252
KEYWORDS mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 11673)
D'Arcangelo,G., Miao,G.G., Chen,S.C., Soares,H.D., Morgan,J.I. and Curran,T.
A protein related to extracellular matrix proteins deleted in the mouse mutant reeler
JOURNAL Nature 374 (6524), 719-723 (1995)

95231649
7715726
2 (bases 1 to 11673)
Direct Submission
Submitted (10-APR-1995) Roche Institute of Molecular Biology, 340
Kingsland St., Nutley, NJ 07110, USA
3 (bases 1 to 11673)
D'Arcangelo,G.
Direct Submission
Submitted (19-DEC-1997) Dev. Neurobiology, St. Jude Childr. Res.
Hosp, 332 N. Lauderdale, Memphis, TN 38105, USA
Sequence update by submitter
On Dec 19, 1997 this sequence version replaced gi:902486.
Location/Qualifiers
1. 11673
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="5"
/map="D5Gmr1"
/tissue_type="cerebellum"
/dev_stage="adult"
1. 282
283..10668
/codon_start=1
/product="reelin"
/protein_id="AAB91599.1"
/db_xref="GI:2702253"
/translation="MERGWAAPRALVLAVLLLLATLARAATGYVPRESFFFLCTHH
GELEGDEGEVLISLHAGNPTYYVPGQYHYHTISTFTFDGLLVTGLTSTSIQSS
QSHGSAFGFMSDHFQNCMSVAVSHVSHLPTNLNFSVNIADPAAGTQOGLNPN
TATHRGQVIFKDALAOQLCEQATATYATSHLAHSDSVILRRDFTASQOGLNPN
IWECSNCEMGEQGTIMHGNATFCPEYPRELTTLTNTTASVLFQSIGSGSFR
SYSDPSITVSYAKNNTADQLEKIRAPSNVSTVHILYLPPEAKGSVQFQWQDNL
RGEVYEACWALDNILVINSAREVLENDLPDVTGNLWLFPGCATVKHSCQSDNSI
YFNGESSEFNATRDVLDSTEDIQWSEEEESQPTGNDILGAVVACGCTVESL
SLVFKDGERKLCPTMTTGTGNLREFFVWVGICDPGVSHENDILYIKEGKEHL
ALDTLAYSYKPSLVSVINPELQTPATKFCRKSHQGNRNWVAFFHVLPLP
LMTYIQFISNLGCGTHQPSVLEFSTNHSRSLHTECLPEACGPHLPSTV
YSGENTSGMNRITIPLNAALTRDTIRWRQGTILGNMWAIDNVYIGPSCLKFCGR
QCTRHCKCDPGFSPACEMASQTFPMFISESGFSARSLSYHNFYSIRGAESVFGG
VLASGALVFNKDGROLITSFIDSQSRELOPTRLGSKSVLSTCAPDQCGEVL
HYSDNGITWKLLEHYVYNHEPRIISVELPDDARQFGIOFRMQPYHSSQGEDVNA
IDEIVMTSLFNSISLDTNLVEVTSGLVGLNVQPCGHDMTLCFTGDSKLASSMR
YVETQSMQIGASYMIQFSLVWGGCQKTPHMDNQVKLEYSAHGLTWHVLQEECLPSM
PSQCEFTSASITHASEFTQWRVTVPKQTSWGTATFRMSQSYTTADSWALDNIYI
GCQPNMCSGHGCDHGVCDQCYQTECHPEALPSTIMSDPENPSWSDMOEVI
GGVWPECGCGVSVSSGLYFSKAGKRLVSMDLDTSWDFVQFYLIQGESNACNK
PDREBGLLIQYNNGGIQHLLAEMFSDFSKPRFVYLELPAAGTKPTCFRFRKKPV
FSGEDYDQWAVDDIIILSEKOKQVIVVNPPTLPQNFYEKPAFDYPMNOMSVMLANE
GMAKNSFCATTPSAMVFGKSDGRFAVTRDLTLKPGYVLFQKLNIGCTSQFSSTAPV
LLQYSHDAGMSWFLLEKGCPPASAAKCGNSRELSPTVYTGDFEWRITITAIIPR
SLASSKTRFRIQESSQKNVPPGLGVYISPCPSYCSGHGDCISGVCFDLGYTA
AQCCTVSNTPNHSEMFDFRFGKLSPLWYKTTGGVCTCGCTLNDGRSLYFNGLKREA
RTVPLDTRNISLVQFYLIQSKTSGITTYITPRARYEGLVYQYSDNDGILWHLLRELDF
MSLEPQIISLDLPREKATFATFRWQPHQKHAQWALDVLIGVNDSSQTGFQDK
LDGSDIDANNYRIQGGVDIDICLMDLTALIFTEINIGNPRIATWDFHVSSEFLQWE
MNMGCKSPFSGAHIQIOLYSINNGKQWLTVECPPTIGCVHTSTSTYTSERFQW
RVTYVLPATNSPRTFRWITQNTYVAGDSWAIDNVILASGCPWMSGRIGICSGRGC
VCDRGGEGCPVPVPLPSILKDDFNGLHPDLWPEVYGAERGNLNGETIKSGTCLIF
KGEGLMLISRDLCNTMYVQPSLRFIAKGTSPERSHILLOFSYSGCVTWHLMDEFF
FPQTTSLFINVPLPYGAQTNATFRWQPHQKHAQWALDVLIGVNDSSQTGFQDK
LDTFDFGPREDNMFYFPGNIGLYCYSSKGAPEEDSAMFVNEVEGHSITTRDLSV
NENTIIQFENVGCSTDSADPVRLEFDFGATWHLPLCYHSSLSVSSLCSTEH
HPSTYVAGTQGRREVHFHGLHLCGSVRFWYQGYFAGSQPVTWADNVYIGQ
CEMCGYHSGCINGTKICDPYSGPTCKISTKNPDFLKDDEGOLSEDFLLMSGCK
PSRCKGLTSGNLLFNEDEGLRMLVTRDLDLSHARFVQFRWLRCGCKGCDPDRSQPVL
LOYSNLGGLSWSLQLEFLFNSNNGVRIALEMPLKARSTRLRWQWQPSNGHFYSP
WVYDQIILIGNISNTVLEDDFTLDSRKLWHPGKTMPCVCGSTGALVFIERAKR
SYVHLORILVSDTFNKWTRITLPLPSYTRSOATFRWHPKAPDFKQOTWADNVYIGD
SGHLCDCSHGRCVQSGCDEOMGLYCDDEPISLTQLKDNFNRAFSNOMLTVSGS
KLSTVCGAVASGLALHFSGGCSRLITVTLNLNNAEFIQFYFMYGCLITPSNRNQGLV

HEATTYNAVNSWKRITIQIDPHVSSATQFRWIOKGETEKOSWAIDHVTGEACP
RLCSGHGYCTGAVCTCDESFQDGSVFSHELPSYIKDNFESARTEANWETIOGGA
IGSCGGLAPYAGHDSLYFNGCQIRQAATKPLDTRASKIMFVLQIGSTAOQSCNSD
LSPGHTVDDKAVLLQYSVNNGITWHVIAQHOPKDFTOAORVSYNVPLEARMKGVLLRW
QPHNTEGHQDOWALDHVEVVLVSTRQYMMNFSRQGLRHFRNRRRSLLRRYP"

BASE COUNT 2645 a 3029 c 2930 g 2583 t
ORIGIN

Alignment Scores:
Pred. No.: 2,25e-62 Length: 11187
Score: 596.00 Matches: 111
Percent Similarity: 97.44% Conservative: 3
Best Local Similarity: 94.87% Mismatches: 3
Query Match: 95.36% Indels: 0
DB: 10 Gaps: 0

US-09-897-438B-2 (1-117) x AB049473 (1-11187)

Qy 1 GluGlnCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20
|||||
Db 1031 GAACAGTGTGGCACCATCATCGATGGCAACGCTGCACCTTCTGTGAGCCATATGGTCT 1090
Qy 21 ArgGluLeuThrThrCysLeuAsnThrThrThrAlaSerValLeuGlnPheSerIle 40
|||||
Db 1091 CGAGATTTGACTACCATACCTGAACACGACACAGCATCCGTCCTCCAGTCTCCATT 1150
Qy 41 GlySerGlySerCysArgPheSerTyrSerAspProSerIleThrValSerTyrAlaLys 60
|||||
Db 1151 GGGTCAGGATCTCGATTTAGTACTCTGACCCAGCATCATTTGTATATGCCAAG 1210
Qy 61 AsnAsnThrAlaAspTrpIleGlnLeuGluLysIleArgAlaProSerAsnValSerThr 80
|||||
Db 1211 AACAACTACTGCGATTTGATTCAGTGGAGAGATTAGAGCCCTTCCAACGTAAGCACC 1270
Qy 81 ValIleHisIleLeuTyrLeuProGluGluAlaLysGlyGluSerValGlnPheGlnTrp 100
|||||
Db 1271 ATCATCCACATPCTCTACCTCTCTGAGAGACGCCAAAGGGAGACGTCGAGTTCACGTGG 1330
Qy 101 LysGlnAspSerLeuArgValGlyGluValTyrGluAlaCysTrpAlaLeu 117
|||||
Db 1331 AACAGCAGCAGCTGCATGTGGTGAGGTGTACGAAGCCTGCTGGGCCCTG 1381

RESULT 4
AX410790
LOCUS AX410790 11580 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 3437 from Patent WO0229103.
ACCESSION AX410790
VERSION AX410790.1 GI:21443495
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 3437 11-APR-2002;
GENE LOGIC INC (US)

FEATURES
source Location/Qualifiers
1..11580
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. U79716"
BASE COUNT 3014 a 2696 c 2753 g 3116 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 2.1e-59 Length: 11580
Score: 572.00 Matches: 106
Percent Similarity: 95.73% Conservative: 6
Best Local Similarity: 90.60% Mismatches: 5
Query Match: 91.52% Indels: 0
DB: 6 Gaps: 0

US-09-897-438B-2 (1-117) x AX410790 (1-11580)

Qy 1 GluGlnCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20
|||||
Db 860 GAACAGTGTGGCGGATATGATGCGCAATGCGCTCACCTTCTGTGAACCATATGGCCCA 919
Qy 21 ArgGluLeuThrThrCysLeuAsnThrThrThrAlaSerValLeuGlnPheSerIle 40
|||||
Db 920 CGAAGCTGATTACACACAGGCCCTTAATACACACACAGCTTCTGCTCCCAATTTCCATT 979
Qy 41 GlySerGlySerCysArgPheSerTyrSerAspProSerIleThrValSerTyrAlaLys 60
|||||
Db 980 GGGTCAGGTCATCGCTTTAGTTATTACAGACCCAGCATCATCGTGTATATGCCAAG 1039
Qy 61 AsnAsnThrAlaAspTrpIleGlnLeuGluLysIleArgAlaProSerAsnValSerThr 80
|||||
Db 1040 AATAACCTCTCGGACTGGATTCAGCTAGAGAAATAGAGCCCTTCCAATGTGAGCACA 1099
Qy 81 ValIleHisIleLeuTyrLeuProGluGluAlaLysGlyGluSerValGlnPheGlnTrp 100
|||||
Db 1100 ATCATCCATATCTCTACCTTCTGAGGACGCCAAAGGGAGATGTCATATTCAGTGG 1159
Qy 101 LysGlnAspSerLeuArgValGlyGluValTyrGluAlaCysTrpAlaLeu 117
|||||
Db 1160 AAGCAGGAAATCTCTGTAGTGAAGTGTATGAAGCCTGCTGGGCTTA 1210

RESULT 5
HSU79716
LOCUS HSU79716 11580 bp mRNA linear PRI 25-FEB-1997
DEFINITION Human reelin (RELN) mRNA, complete cds.
ACCESSION U79716
VERSION U79716.1 GI:1743884
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 11580)
AUTHORS Desilva,U., D'Arcangelo,G., Braden,V.V., Chen,J., Miao,G.G.,
Curran,T. and Green,E.D.
TITLE The human reelin gene: isolation, sequencing, and mapping on
chromosome 7
JOURNAL Genome Res. 7 (2), 157-164 (1997)
MEDLINE 97202106
PUBMED 9049633
REFERENCE 2 (bases 1 to 11580)
AUTHORS Desilva,U., D'Arcangelo,G., Braden,V.V., Chen,J., Miao,G.G.,
Curran,T. and Green,E.D.
TITLE Direct Submission
JOURNAL Submitted (26-NOV-1996), National Center for Human Genome Research,
National Institutes of Health, 49 Convent Drive, MSC4431, Bethesda,
MD 20892, USA

FEATURES
source Location/Qualifiers
1..11580
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
1..11580
/gene="RELN"
176..10358
/gene="RELN"
/codon_start=1
/product="reelin"
/protein_id="AAC51105.1"
/db_xref="GI:1743885"
/translation="MERSGWARTFLALLGLATLRAARAAGYPRFSPFFLCTHHG
ELEGGDEQGEVLISLHIAAGNTYVPGQYHVTISTFDFGLLVGLYTSVQASQ
SIGGSAIFGDMDSHQFNGFVSVASHSLPTNLSFIWIAPAGTGCYVNEAT
ATHRGQVIFKDALAQOLCEQAPDVTVHPHLAEIHSDSILRLDDFDSYHOLQNPNI
WYECNCTGEQCCGATMHGNNAVTFCEPYGPRELITGLNTTASVLOFISGSSCRFS
YSDPSILVIYAKNNSADWOLEKIRAPSNVSTIIHILYLPEDAGKENVQPMKQENLR
VGEVVEACWALDNILITNSAHRQVLEDSLDPDVDTGNWULFFPGATVKHSCQSDGNSIY

pH: 8.6

This STS was developed from sequence determined by another investigator. See GenBank record: U79716 For additional information about the NHGRI chromosome 7 mapping project, see <http://www.nhgri.nih.gov/DIR/GB/CHR7>. Also see Genomics 11:548-64 (1991) [MUID-92128937].

FEATURES

source
1. .11580
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="7"
/clone_lib="Eric D. Green"
1. .11580
/gene="RELN"
1112. 1294
/gene="RELN"
1112. 1134
/gene="RELN"
primer_bind
primer_bind complement(1272. 1294)
BASE COUNT 3014 a 2696 c 2753 g 3116 t 1 others
ORIGIN

gene

STS

primer_bind

Alignment Scores:
Pred. No.: 2,1e-59 Length: 11580
Score: 572.00 Matches: 106
Percent Similarity: 95.73% Conservative: 6
Best Local Similarity: 90.60% Mismatches: 5
Query Match: 91.52% Indels: 0
DB: 11 Gaps: 0

US-09-897-438B-2 (1-117) x G30936 (1-11580)

Qy 1 GluGlnCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20
Db 860 GAACAGTGTGGCGGATTCATGTCGCAATGCCCTCTGTGAACCATATGGCCCA 919
Qy 21 ArgGluLeuThrThrCysLeuAsnThrThrAlaSerValLeuGlnPheSerIle 40
Db 920 CGAGAACTGATTACCAAGCGCTTAATACACACAGCTTCTGTCCTCCATTTCCATT 979
Qy 41 GlySerGlySerCysArgPheSerTyrSerAspProSerIleThrValSerTyrAlaLys 60
Db 980 GGGTCAGGTTTCATGTCGCTTTAGTTATTTCAGACCCCAAGCATCATCGTGTATGCCAAG 1039
Qy 61 AsnAsnThrAlaAspTrpIleGlnLeuGluLysIleArgAlaProSerAsnValSerThr 80
Db 1040 AATACCTCGCGGACTGGATTACGTAGAGAAATFAGAGCCCTTCCAAATGTCAGCACA 1099
Qy 81 ValIleHisIleLeuTyrLeuProGluGluAlaLysGlyGluSerValGlnPheGlnTrp 100
Db 1100 ATCATCCATATCCTCTACCTTCTGAGGACGCCAAAGGGAGAAATGTCCAATTTTCAGTGG 1159
Qy 101 LysGlnAspSerLeuArgValGlyGluValTyrGluAlaCysTrpAlaLeu 117
Db 1160 AAGCAGGAAAAATCTCGTGTAGTGAGTGATGATGAAGCCTCTGGGCCTTA 1210

RESULT 7

G30938 G30938 11580 bp DNA linear STS 28-SEP-1998
LOCUS SWSS3176 Eric D. Green Homo sapiens STS genomic, sequence tagged
DEFINITION site.

ACCESSION G30938

VERSION G30938.1 GI:1923211

KEYWORDS STS.

SOURCE Homo sapiens.

ORGANISM Homo sapiens

REFERENCE
1 (bases 1 to 11580)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Bouffard,G.G., Iyer,L.M., Idol,J.R., Braden,V.V., Cunningham,A.F.,
Weintraub,L.A., Mohr-Tidwell,R.M., Peluso,D.C., Fulton,R.S.,
Leckie,M.P. and Green,E.D.

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

A collection of 1814 human chromosome 7-specific STS
Genome Res. 7 (1), 59-64 (1997)
97189344
9037602
2 (bases 1 to 11580)
Green,E.D.
Human chromosome 7 STS (1997)
Unpublished (1997)
On Apr 3, 1997 this sequence version replaced gi:1706937.
Synonyms: RELN
GDB_DSEG: RELN
Contact: Eric D. Green
Genome Technology Branch
National Human Genome Research Institute/NIH
49 Convent Dr., MSC4431, Bldg. 49, Rm. 2A08, Bethesda, MD 20892
Tel: 3014020201
Fax: 3014024735
Email: egreen@nhgri.nih.gov
Primer A: TGTGGCTGATTTACTAACAG
Primer B: TGGGCTTTTATAGCTTG
STS size: 238
PCR Profile:

Presoak: 0 degrees C for 0.00 minute(s)
Denaturation: 92 degrees C for 0.17 minute(s)
Annealing: 50 degrees C for 1.00 minute(s)
Polymerization: 72 degrees C for 1.00 minute(s)
PCR Cycles: 35
Thermal Cycler: PerkinElmer 9600

Protocol:
Template: 30-100 ng
Primer: each 1 uM
dNTPs: each 200 uM
Taq Polymerase: 0.05 units/ul
Total Vol: 10 ul

Buffer:
MgCl2: 1.5 mM
KCl: 100 mM
Tris-HCl: 10 mM
NH4Cl: 5 mM
pH: 8.6

This STS was developed from sequence determined by another investigator. See GenBank record: U79716 For additional information about the NHGRI chromosome 7 mapping project, see <http://www.nhgri.nih.gov/DIR/GB/CHR7>. Also see Genomics 11:548-64 (1991) [MUID-92128937].

FEATURES

source
1. .11580
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="7"
/clone_lib="Eric D. Green"
1. .11580
/gene="RELN"
11314. 11551
/gene="RELN"
11314. 11333
/gene="RELN"
primer_bind
primer_bind complement(11535. 11551)
BASE COUNT 3014 a 2696 c 2753 g 3116 t 1 others
ORIGIN

Alignment Scores:

Pred. No.: 2,1e-59 Length: 11580
Score: 572.00 Matches: 106
Percent Similarity: 95.73% Conservative: 6
Best Local Similarity: 90.60% Mismatches: 5
Query Match: 91.52% Indels: 0
DB: 11 Gaps: 0

US-09-897-438B-2 (1-117) x G30938 (1-11580)


```

RESULT 9
AC128022/c
LOCUS
DEFINITION
Rattus norvegicus clone CH230-525L20, *** SEQUENCING IN PROGRESS
***, 97 unordered pieces.

ACCESSION
AC128022
VERSION
AC128022.1 GI:21908605
KEYWORDS
HTG; HTGS_PHASE1.
SOURCE
Rattus norvegicus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 202764)
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,
Alsbrooks, S.L., Amarantunge, H.C., Are, J.R., Ayele, M., Banks, T.,
Barbaria, J., Benton, J., Blmage, K., Blankenburg, K., Bonnin, D.,
Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carroll, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Haviak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,
Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karlssoen, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtharge, O., Lieu, C., Liu, J., Liu, W., Louissegh, H.,
Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mahoney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenwo, S., Oguh, M., Okwuonu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Picken, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I.,
Sodergren, E., Sonake, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, Y., Villalon, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 202764)
Worley, K.C.
Direct Submission
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KAIV
Center clone name: CH230-525L20
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye: 100% of reads

```

Assembly program: Phrap; version 0.990329
Consensus quality: 113069 bases at least Q40
Consensus quality: 119970 bases at least Q30
Consensus quality: 125656 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 97 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1244: contig of 1244 bp in length
* 1245: gap of unknown length
* 1345: contig of 1002 bp in length
* 2346: gap of unknown length
* 2347: contig of 1367 bp in length
* 2447: gap of unknown length
* 3814: contig of 1188 bp in length
* 3914: gap of unknown length
* 5102: contig of 1523 bp in length
* 5202: gap of unknown length
* 6725: contig of 1224 bp in length
* 6825: gap of unknown length
* 8049: contig of 1064 bp in length
* 8149: gap of unknown length
* 9212: contig of 1120 bp in length
* 9312: gap of unknown length
* 9313: contig of 1358 bp in length
* 10433: gap of unknown length
* 10533: contig of 1078 bp in length
* 11891: gap of unknown length
* 11990: contig of 1318 bp in length
* 13068: gap of unknown length
* 13169: contig of 1410 bp in length
* 14487: gap of unknown length
* 14596: contig of 1438 bp in length
* 15997: gap of unknown length
* 16096: contig of 1069 bp in length
* 16097: gap of unknown length
* 17534: contig of 1081 bp in length
* 17535: gap of unknown length
* 17635: contig of 1555 bp in length
* 17636: gap of unknown length
* 18704: contig of 1518 bp in length
* 18803: gap of unknown length
* 18804: contig of 1255 bp in length
* 19884: gap of unknown length
* 19885: contig of 1238 bp in length
* 19985: gap of unknown length
* 21540: contig of 1421 bp in length
* 21541: gap of unknown length
* 21640: contig of 1180 bp in length
* 22781: gap of unknown length
* 22881: contig of 1041 bp in length
* 23927: gap of unknown length
* 24027: contig of 1517 bp in length
* 25545: gap of unknown length
* 25546: contig of 1484 bp in length
* 25645: gap of unknown length
* 26900: contig of 1552 bp in length
* 27000: gap of unknown length
* 28238: contig of 1282 bp in length
* 28338: gap of unknown length
* 28339: contig of 1434 bp in length
* 29759: gap of unknown length
* 29859: contig of 1434 bp in length
* 31039: gap of unknown length
* 31139: contig of 1434 bp in length
* 32180: gap of unknown length
* 32280: contig of 1517 bp in length
* 33797: gap of unknown length
* 33897: contig of 1484 bp in length
* 35381: gap of unknown length
* 35481: contig of 1552 bp in length
* 37033: gap of unknown length
* 37133: contig of 1282 bp in length
* 38415: gap of unknown length
* 38515: contig of 1434 bp in length
* 39949: gap of unknown length

Db 15439 TTCACACACGGCTGTGGTCTGAAC¹TGTGAGC²TTGACCTGCCTTCTCCCAC³TTGTTCC⁴---

```

Qy 57 SerTyrAlaLysAsnThrAlaAspTrpIleGlnLeuGluLysIleArgAlaProSer 76
Db 15382 -----AGAGCCCTTCC 15371

Qy 77 AsnValSerThrValIleHisIleLeuTyrLeuProGluAlaLysGlyGluSerVal 96
Db 15370 AATGTGACACAGTCATCCATCTGTACCTCCCGAGGAAGCAAGGGAGAGCGTG 15311

Qy 97 GlnPheGlnTrpLysGlnAspSerLeuArgValGlyGluValTyrGluAlaCysTrpAla 116
Db 15310 CAGTTCAGTGAACAGGACGCTCGAGTGGGTATGAGGCTGCTGGGCC 15251

Qy 117 Leu 117
Db 15250 CTG 15248

RESULT 11
LOCUS AC023062 183156 bp DNA linear HTG 05-MAY-2000
DEFINITION Mus musculus chromosome 5 clone RP23-466D2 strain C57BL6/J, WORKING
DRAFT SEQUENCE, 8 unordered pieces.
AC023062
AC023062.2 GI:7709916
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Mus musculus.
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 183156)
Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W.,
Bouffard,G.G., Dietrich,N.L., Eagle,W.O., Gupta,J., Ho,S.-L.,
Huang,M.C., Idol,J., Lee-Lin,S.-O., Maduro,Q.L., Maduro,V.B.,
Mastrian,S.D., McCloskey,J.C., Morse,E., Ojodu,M.A., Pearson,R.,
Stantripop,S., Summers,T.J., Thomas,J.W., Thomas,P.J.,
Tlionson,E.E., Touchman,J.W., Tran,J.T., Vogt,J.L., Walker,M.A.,
Wetherby,K.D. and Green,E.D.
NISC Mouse Sequencing Initiative
Unpublished
2 (bases 1 to 183156)
Green,E.D.
Direct Submission
Submitted (08-FEB-2000) NIH Intramural Sequencing Center, 8717
GroveMont Circle, Gaithersburg, MD 20877, USA
On May 5, 2000 this sequence version replaced gi:6939141.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.mouse@nhgri.nih.gov
----- Project Information
Center project name: y1
Center clone name: 466D02
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 175371 bases at least Q40
Consensus quality: 178474 bases at least Q30
Consensus quality: 180133 bases at least Q20
Insert size: 179000; agarose-fp
Insert size: 188000; pulse-field-gel
Insert size: 183156; sum-of-contigs
Quality coverage: 6.96x in Q20 bases; agarose-fp
Quality coverage: 6.63x in Q20 bases; pulse-field-gel
Quality coverage: 6.80x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence

```

* as soon as it is available and the accession number will be preserved.

1 2374: contig of 2374 bp in length

* 2375 4490: gap of unknown length

* 2375 4490: contig of 2116 bp in length

* 4491 14959: gap of unknown length

* 4491 14959: contig of 10469 bp in length

* 14960 31072: gap of unknown length

* 31073 50119: contig of 16113 bp in length

* 50120 78196: gap of unknown length

* 78197 109874: contig of 19047 bp in length

* 109875 183156: gap of unknown length

* 109875 183156: contig of 28077 bp in length

* 109875 183156: contig of 31678 bp in length

* 109875 183156: contig of 73282 bp in length.

FEATURES

source

1. .183156

/organism="Mus musculus"

/strain="C57BL6/J"

/db_xref="taxon:10090"

/chromosome="5"

/clone="RP23-466D2"

/clone_lib="RPCI mouse BAC library 23"

BASE COUNT 53719 a 37080 c 37302 g 55052 t 3 others

ORIGIN

Alignment Scores:

Pred. No.: 4.84e-17 Length: 183156

Score: 240.00 Matches: 50

Percent Similarity: 66.67% Conservative: 4

Best Local Similarity: 61.73% Mismatches: 9

Query Match: 38.40% Indels: 18

DB: 2 Gaps: 2

US-09-897-438B-2 (1-117) x AC023062 (1-183156)

Qy 38 PheSerIleGlySerGly---SerCysArgPheSerTyrSerAspProSerIleThrVal 56

Db 173566 TTCACACACGGCGTGGTCTGAACTGTGAGCTTGACCTGCCTCTCCCATCTGTTC--- 173622

Qy 57 SerTyrAlaLysAsnThrAlaAspTrpIleGlnLeuGluLysIleArgAlaProSer 76

Db 173623 -----ACAGCCCTTCC 173634

Qy 77 AsnValSerThrValIleHisIleLeuTyrLeuProGluAlaLysGlyGluSerVal 96

Db 173635 AATGTGACACAGTCATCCATCTGTACCTCCCGAGGAAGCAAGGGAGAGCGTG 173694

Qy 97 GlnPheGlnTrpLysGlnAspSerLeuArgValGlyGluValTyrGluAlaCysTrpAla 116

Db 173695 CAGTTCAGTGAACAGGACGCTCGAGTGGGTATGAGGCTGCTGGGCC 173754

Qy 117 Leu 117

Db 173755 CTG 173757

RESULT 12

LOCUS AC124933 183641 bp DNA linear HTG 31-JUL-2002

DEFINITION Rattus norvegicus clone CH230-466N17, *** SEQUENCING IN PROGRESS

AC024933

AC124933

AC124933.2 GI:21952716

VERSION HTG; HTGS_PHASE1.

KEYWORDS Norway rat.

SOURCE Rattus norvegicus

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE 1 (bases 1 to 183641)

AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,

Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T.,
 Barbara, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D.,
 Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
 Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
 Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
 Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
 Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, K.H.,
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
 Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
 Fallis, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodgson, A., Hoques, M., Holloway, C., Hollins, B.,
 Homsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudan, S.,
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H.,
 Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
 Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokenwo, S., Oguh, M., Okwuonu, G.,
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
 Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G.,
 Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I.,
 Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tanssey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Usmani, K., Vasquez, L., Vera, V., Villalobos, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstock, G. and Gibbs, R.

Direct Submission
 Unpublished
 2 (bases 1 to 183641)
 Worley, K.C.

Direct Submission
 Submitted (20-JUN-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 183641)
 Worley, K.C.

Direct Submission
 Submitted (31-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 24, 2002 this sequence version replaced gi:21490053.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: KAVI
 Center clone name: CH230-466N17
 ----- Summary Statistics
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 111724 bases at least Q40
 Consensus quality: 120970 bases at least Q30
 Consensus quality: 128458 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently

* consists of 88 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 1107: contig of 1107 bp in length
 * 1108: gap of unknown length
 * 1207: contig of 1207 bp in length
 * 1208: contig of 1253 bp in length
 * 2461: gap of unknown length
 * 2561: contig of 1444 bp in length
 * 4004: gap of unknown length
 * 4104: gap of unknown length
 * 5382: contig of 1278 bp in length
 * 5383: gap of unknown length
 * 5482: contig of 1392 bp in length
 * 5483: gap of unknown length
 * 6874: gap of unknown length
 * 6875: contig of 1097 bp in length
 * 8071: contig of 1097 bp in length
 * 8171: gap of unknown length
 * 8172: contig of 1307 bp in length
 * 9478: gap of unknown length
 * 9479: gap of unknown length
 * 9578: gap of unknown length
 * 9579: contig of 1483 bp in length
 * 11061: gap of unknown length
 * 11161: contig of 1581 bp in length
 * 11162: gap of unknown length
 * 12742: gap of unknown length
 * 12743: gap of unknown length
 * 12842: gap of unknown length
 * 12843: contig of 1662 bp in length
 * 14504: gap of unknown length
 * 14505: gap of unknown length
 * 14506: gap of unknown length
 * 15734: contig of 1130 bp in length
 * 15834: gap of unknown length
 * 15835: contig of 1124 bp in length
 * 16958: gap of unknown length
 * 17058: gap of unknown length
 * 17059: contig of 1646 bp in length
 * 18704: gap of unknown length
 * 18705: gap of unknown length
 * 20420: contig of 1616 bp in length
 * 20520: gap of unknown length
 * 21648: contig of 1128 bp in length
 * 21649: gap of unknown length
 * 23177: contig of 1429 bp in length
 * 23178: gap of unknown length
 * 23278: contig of 1483 bp in length
 * 24760: gap of unknown length
 * 24761: gap of unknown length
 * 24861: contig of 1899 bp in length
 * 26760: gap of unknown length
 * 26859: gap of unknown length
 * 28299: contig of 1440 bp in length
 * 28300: gap of unknown length
 * 28301: gap of unknown length
 * 29592: contig of 1193 bp in length
 * 29593: gap of unknown length
 * 30870: contig of 1178 bp in length
 * 30871: gap of unknown length
 * 32642: contig of 1672 bp in length
 * 32643: gap of unknown length
 * 32742: gap of unknown length
 * 34386: contig of 1644 bp in length
 * 34486: gap of unknown length
 * 36064: contig of 1578 bp in length
 * 36164: gap of unknown length
 * 37398: contig of 1234 bp in length
 * 37399: gap of unknown length
 * 37499: contig of 2110 bp in length
 * 39608: gap of unknown length
 * 39708: contig of 1772 bp in length
 * 41480: gap of unknown length
 * 41481: contig of 1179 bp in length
 * 42759: gap of unknown length
 * 42760: gap of unknown length
 * 42860: contig of 1852 bp in length
 * 44711: gap of unknown length
 * 44811: contig of 1397 bp in length
 * 46208: gap of unknown length
 * 46308: gap of unknown length
 * 46309: contig of 1459 bp in length
 * 47667: gap of unknown length
 * 47668: gap of unknown length
 * 49227: contig of 1360 bp in length
 * 49228: gap of unknown length
 * 49327: contig of 1470 bp in length
 * 50797: gap of unknown length
 * 50798: gap of unknown length


```

Qy 93 GlyGluSerValGlnPheGlnTrpLysGlnAspSerLeuArgValGlyGluValTyrGlu 112
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 92887 GGGGAGAACGTCAGTCCAGTGAACAGACAGACGCTGCTGGGTGAGGTGTACGAA 92946

Qy 113 AlaCysTrpAlaLeu 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 92947 GCCTGCTGGGCCCTG 92961

RESULT 14
HSAC000121/c 93163 bp DNA linear PRI 31-JAN-1997
LOCUS Human BAC clone RG249A12 from 7q22, complete sequence.
DEFINITION AC000121
ACCESSION AC000121
VERSION AC000121.1 GI:1809222
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 93163)
AUTHORS Lamar,B., Wamsley,P. and Gibson,A.
TITLE The sequence of H. sapiens BAC clone RG249A12
JOURNAL Unpublished (1997)
REFERENCE 2 (bases 1 to 93163)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-1997)
COMMENT Department of Genetics, Washington University
St. Louis, MO 63108, USA
http://genome.wustl.edu/gsc
e-mail: sapiens@watson.wustl.edu

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
sections once, or longer because we provide a small overlap between
neighboring submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded or sequenced with an alternate
chemistry; an attempt was made to resolve all sequencing problems,
such as compressions and repeats; all regions were covered by
sequence from more than one subclone; and the assembly was
confirmed by restriction digest.

SOURCE INFORMATION:
This clone is from a release of the human BAC library. The library
contains cloned DNA from a human sperm. For references see:
Shizuya et al., Proc. Natl. Acad. Sci. 89:8794-8797 (1992); Kim et
al., Genomics 34:213-218 (1996).
VECTOR: pBEO
Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the right is H_RG352J05. Actual start of
this clone is at base position 1 of H_RG249A12; actual end is at
93163 of H_RG249A12. This clone is part of an unanchored island,
orientation is unknown.

This clone contains STS SWSS2926 (NID:gl706935).
FEATURES
source
1. .93163
Location/Qualifiers
/db_xref="taxon:9606"
/organism="Homo sapiens"
/chromosome="7"
/map="7q22"
/clone="RG249A12"
/clone_lib="CITB-HS-A"
complement(1011..1309)
repeat_region
rpt_family="ALU"
repeat_region 2887..3178
rpt_family="ALU"

```

```

misc_feature 5965..6163
/notes="similar to C2H2-type zinc fingers PROSITE; PS00028"
repeat_region complement(7118..7170)
rpt_family="L1"
repeat_region 9584..12339
rpt_family="L1"
exon 10207..11298
/notes="GRAIL prediction, score = 85"
evidence=not_experimental
complement(10583..11001)
rpt_family="L1"
repeat_region 12423..12714
rpt_family="ALU"
repeat_region 12715..12734
rpt_family="L1"
repeat_region 13365..13412
rpt_family="L1"
repeat_region 14286..14309
rpt_family="L1"
repeat_region 15942..15963
rpt_family="L1"
gene complement(<17055..>84441)
/genes="WUGSC:H_RG249A12.1"
CDS complement(Join(<17055..17200,32792..33032,35849..35945,
58079..58130,63050..63146,84365..>84441))
/genes="WUGSC:H_RG249A12.1"
/notes="extracellular matrix protein; 97% similarity to
mouse reelin U24703 (PID:g902487), partial CDS"
/codon_start=1
evidence=not_experimental
/protein_id="AAB46357.1"
/db_xref="GI:1809223"
translation="EIHSDSITILRDFDSYHOLQNPNIWECNCCETGCGCAIMHG
NAVTFEPYGPRLITGLTNTTASVLOFSGSCRFYSYSDPSIIVLYAKNSADMI
QLEKTRAPSNVSTIHLXLPEDAKGENVOFQKQENLRVGEVYECACWALDNLINS
AIRQVVLSDLPDVTGNNLFFPGATVKHSCSDGNSIYFHGNEGSEFNATTRDVL
STEDIQEQWSEEFESQPTG"
exon complement(17055..17252)
/genes="WUGSC:H_RG249A12.1"
/notes="GRAIL prediction, score = 100"
evidence=not_experimental
21484..21521
rpt_family="L1"
repeat_region complement(21962..21990)
rpt_family="L1"
repeat_region 24430..24468
rpt_family="L1"
repeat_region 24613..24654
rpt_family="L1"
repeat_region 25141..25180
rpt_family="L1"
repeat_region 26470..26761
rpt_family="ALU"
misc_feature complement(26762..26936)
/genes="WUGSC:H_RG249A12.1"
repeat_region 26937..26989
/notes="match to Human cDNA, 5' end H56746 (NID:1005390)"
exon 27107..27238
/notes="GRAIL prediction, score = 82"
evidence=not_experimental
27993..28465
rpt_family="L1"
repeat_region 28476..28496
rpt_family="L1"
repeat_region complement(29663..29903)
rpt_family="ALU"
exon complement(32793..33032)
/genes="WUGSC:H_RG249A12.1"
/notes="GRAIL prediction, score = 90"
evidence=not_experimental
complement(33462..33752)
rpt_family="ALU"

```



```

repeat_region 37975..37994
/rpt_family="L1"
repeat_region 40278..40311
/rpt_family="L1"
repeat_region complement(41921..42075)
/rpt_family="ALU"
repeat_region complement(42447..42734)
/rpt_family="ALU"
exon 43126..43185
/notes="GRAIL prediction, score = 95"
/evidence-not_experimental
repeat_region complement(43411..43680)
/rpt_family="ALU"
repeat_region 45704..45991
/rpt_family="ALU"
misc_feature 46064..46434
/notes="match to Human cDNA H60097, 5' end (NID:g1012929)"
repeat_region complement(46604..46897)
/rpt_family="ALU"
misc_feature complement(47023..47412)
/genes="WUGSC:H.RG249A12.1"
/notes="match to Human cDNA, 3' end H59444 (NID:g1012276)"
repeat_region 58125..58163
/rpt_family="L1"
repeat_region 58604..58758
/rpt_family="L1"
repeat_region 59743..59765
/rpt_family="L1"
repeat_region 60210..60516
/rpt_family="ALU"
exon complement(63051..63146)
/genes="WUGSC:H.RG249A12.1"
/notes="GRAIL prediction, score = 100"
/evidence-not_experimental
repeat_region 64242..64284
/rpt_family="L1"
repeat_region 64445..64496
/rpt_family="L1"
repeat_region complement(67753..68011)
/rpt_family="L1"
repeat_region 67986..68889
/rpt_family="L1"
exon complement(70725..70841)
/genes="WUGSC:H.RG249A12.1"
/notes="GRAIL prediction, score = 82"
/evidence-not_experimental
repeat_region 71911..71947
/rpt_family="L1"
repeat_region complement(73631..73919)
/rpt_family="ALU"
repeat_region 75073..75092
/rpt_family="L1"
repeat_region 76743..82366
/rpt_family="L1"
exon 78326..79402
/notes="GRAIL prediction, score = 83"
/evidence-not_experimental
exon 80586..81548
/genes="GRAIL prediction, score = 82"
/evidence-not_experimental
repeat_region complement(80962..81380)
/rpt_family="L1"
repeat_region complement(82581..82608)
/rpt_family="L1"
misc_feature 83877..84447

Alignment Scores:
Pred. No.: 1,91e-15 Length: 93163
Score: 224.00 Matches: 40
Percent Similarity: 100.00% Conservatve: 5
Best Local Similarity: 88.89% Mismatches: 0
Query Watch: 35.84% Indels: 0
DB: 9 Gaps: 0

```

```

US-09-897-438B-2 (1-117) x HSAC000121 (1-93163)
QY 73 ArgAlaProSerAsnValSerThrValIleHisIleLeuTyLeuProGluGluAlaLys 92
|||||
Db 33034 AGAGCCCTTCCAAATGTCAGCACAAATCATCATCTCTACCTCTCTGAGGAGCCAAA 32975
|||||
QY 93 GlyGluSerValGlnPheGlnTrpLysGlnAspSerLeuArgValGlyGluValTyTyrGlu 112
|||||
Db 32974 GGGGAGAAATGCCAAATTCAGTGGAGCAGGAAATCTTCGTGTAGGTGAAGTGTATGAA 32915
|||||
QY 113 AlaCystTrpAlaLeu 117
|||||
Db 32914 GCCTGCTGGGCCTTA 32900

RESULT 15
AC041023/c
LOCUS
DEFINITION Homo sapiens clone RP11-356G3, WORKING DRAFT SEQUENCE, 20 unordered
pieces.
ACCESSION AC041023
VERSION AC041023.2 GI:8016758
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 185996)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Brown,A., Burkett,G.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Buckle,T., Lehoczy,J.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Spencer,B.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spence,J.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,W.
Direct Submission
Submitted (11-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 22, 2000 this sequence version replaced gi:7534225.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L9601
Center clone name: 356.G.3
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 175155 bases at least Q40

```

Consensus quality: 180623 bases at least Q30
Consensus quality: 182712 bases at least Q20
Insert size: 188000; agarose-fp
Insert size: 184096; sum-of-contigs
Quality coverage: 4.5 in Q20 bases; agarose-fp
Quality coverage: 4.6 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

- * 1 2941: contig of 2941 bp in length
- * 2942 3041: gap of 100 bp
- * 3042 5193: contig of 2152 bp in length
- * 5194 5293: gap of 100 bp
- * 5294 7050: contig of 1757 bp in length
- * 7051 7150: gap of 100 bp
- * 7151 10452: contig of 3302 bp in length
- * 10453 10552: gap of 100 bp
- * 10553 13859: contig of 3307 bp in length
- * 13860 13959: gap of 100 bp
- * 13960 17254: contig of 3295 bp in length
- * 17255 17354: gap of 100 bp
- * 17355 20523: contig of 3169 bp in length
- * 20524 20623: gap of 100 bp
- * 20624 25148: contig of 4525 bp in length
- * 25149 25248: gap of 100 bp
- * 25249 29676: contig of 4428 bp in length
- * 29677 29776: gap of 100 bp
- * 29777 34077: contig of 4301 bp in length
- * 34078 34177: gap of 100 bp
- * 34178 39009: contig of 4832 bp in length
- * 39010 39109: gap of 100 bp
- * 39110 45564: contig of 6455 bp in length
- * 45565 45664: gap of 100 bp
- * 45665 54035: contig of 8371 bp in length
- * 54036 54135: gap of 100 bp
- * 54136 63394: contig of 9259 bp in length
- * 63395 63494: gap of 100 bp
- * 63495 74946: contig of 11452 bp in length
- * 74947 75046: gap of 100 bp
- * 75047 88353: contig of 13307 bp in length
- * 88354 88453: gap of 100 bp
- * 88454 104853: contig of 16400 bp in length
- * 104854 104953: gap of 100 bp
- * 104954 123339: contig of 18386 bp in length
- * 123340 123439: gap of 100 bp
- * 123440 154730: contig of 31291 bp in length
- * 154731 154830: gap of 100 bp
- * 154831 185996: contig of 31166 bp in length.

FEATURES
source

- 1. .185996
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="rp11-356G3"
/location="RP11-11 Human Male BAC"
- 1. .2941
/note="assembly_fragment"
3042. .5193
/note="assembly_fragment"
5294. .7050
/note="assembly_fragment"
7151. .10452
/note="assembly_fragment"
10553. .13859
/note="assembly_fragment"
13960. .17254
/note="assembly_fragment"
17355. .20523
- misc_feature
/note="assembly_fragment"
3042. .5193
/note="assembly_fragment"
5294. .7050
/note="assembly_fragment"
7151. .10452
/note="assembly_fragment"
10553. .13859
/note="assembly_fragment"
13960. .17254
/note="assembly_fragment"

- misc_feature
/note="assembly_fragment"
20624. .25148
/note="assembly_fragment"
25249. .29676
/note="assembly_fragment"
clone_end:T7
vector_side:right"
29777. .34077
/note="assembly_fragment"
34178. .39009
/note="assembly_fragment"
39110. .45564
/note="assembly_fragment"
45665. .54035
/note="assembly_fragment"
54136. .63394
/note="assembly_fragment"
63495. .74946
/note="assembly_fragment"
75047. .88353
/note="assembly_fragment"
88454. .104853
/note="assembly_fragment"
104954. .123339
/note="assembly_fragment"
123440. .154730
/note="assembly_fragment"
154831. .185996
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
BASE COUNT 60145 a 34255 c 33646 g 56049 t 1901 others
ORIGIN

Alignment Scores:

Pred. No.: 3.18e-05 Length: 185996
Score: 144.00 Matches: 36
Percent Similarity: 47.17% Conservative: 14
Best Local Similarity: 33.96% Mismatches: 24
Query Match: 23.04% Indels: 32
DB: 2 Gaps: 4

US-09-897-438b-2 (1-117) x AC041023 (1-185996)

- QY 1 GluGlnCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20
|||||
Db 166372 GAACAGTGTGGCGGATATGTCATGGCAATGCCGTACCTCTGTGAACCATATGGCCCA 166313
QY 21 ArgGluLeuThrThrThrCysLeuAsnThrThrAlaSerValLeuGlnPheSerIle 40
|||||
Db 166312 CGAGAACTGCTAAGTATGCTATTCATGTACA-----TTATTCCATTATAGTATT 166262
QY 41 GlySerGlySerCysArgPheSerTyr-----SerAspProSerIle 54
|||||
Db 166261 TACTGAGCTTTTGGCCCAACAAAAGTTTCTTATGCTCTTAAAGTTCAGATGAAAAT--- 166205
QY 55 ThrValSerTyrAlaLysAsnAsnThrAlaAspTrpIleGlnLeuGluLysIleArgAla 74
Db 166205 ----- 166205
QY 75 ProSerAsnValSerThrValIleHisIleLeuTyrLeuProGluAlaLysGlyGlu 94
|||||
Db 166204 ---TCAAAATATGATTTTATGAGTCTCTATTACAAATATCTT---AAAAATTGGAGGGGTGAT 166151
QY 95 ServAlGlnPheGlnTrp 100
|||||
Db 166150 TCTATGGCCTATTTTGG 166133

Search completed: November 6, 2002, 20:18:34
Job time : 2907 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 6, 2002, 19:24:47 ; Search time 67 Seconds
(without alignments)
535.540 Million cell updates/sec

Title: US-09-897-438B-2

Perfect score: 625

Sequence: 1 EQCGTGHGNAVTFCEPYCP.....FOWKODSLRGEVYACWAL 117

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Fgapop 6.0, Fgapext 7.0

Delopt 6.0, Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV-xlp

-MODEL-frame+p2n.model -DEV-xlp
-Q=/cgn2_1/USPTO.spool/US09897438/runat_06112002_101925_11313/app_query.fasta_1.263
-DB-Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOFC=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09897438 @CGN_1_1_31 @runat_06112002_101925_11313 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	625	100.0	11673	4	US-09-334-220-3
2	572	91.5	11580	4	US-09-334-220-4
3	81	13.0	2802	4	US-09-351-814-12
4	78	12.5	414	4	US-08-905-223-151
5	72.5	11.6	3102	4	US-09-336-643A-17
6	72.5	11.6	28473	4	US-08-961-527-83
7	71	11.4	686	2	US-08-627-151A-12
8	68	10.9	331	4	US-09-060-756-308
9	68	10.9	2231	4	US-08-961-527-265
10	67.5	10.8	2937	6	5208144-7
11	67	10.7	741	4	US-09-134-001C-2455
12	66.5	10.6	2909	4	US-08-104-158-1

ALIGNMENTS

RESULT 1
US-09-334-220-3
; Sequence 3, Application US/09334220
; Patent No. 6323177
; GENERAL INFORMATION:
; APPLICANT: St. Jude's Children's Research Hospital
; APPLICANT: Curran, Thomas
; APPLICANT: D'Arcangelo, Gabriella
; TITLE OF INVENTION: INTERACTION OF REELIN WITH VERY LOW
; TITLE OF INVENTION: DENSITY LIPOPROTEIN (VLDL) RECEPTOR FOR SCREENING AND
; TITLE OF INVENTION: THERAPIES
; FILE REFERENCE: 2427/0F704
; CURRENT APPLICATION NUMBER: US/09/334,220
; CURRENT FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 11673
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-334-220-3

Alignment Scores:
Pred. No.: 4.64e-76 Length: 11673
Score: 625.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-897-438B-2 (1-117) x US-09-334-220-3 (1-11673)

QY 1 GUGlncCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20
|||||
Db 970 GAGCAGTGTGGCCACCATCATGCATGCATGCTACCTTCTGTGAGCGCGTACGCGCT 1029

c	13	66	10.6	318	4	US-09-240-274-214	Sequence 214, App
	14	66	10.6	1509	2	US-08-791-347-2	Sequence 2, Appli
	15	66	10.6	1985	2	US-08-791-347-1	Sequence 1, Appli
	16	66	10.6	3294	1	US-08-409-995-1	Sequence 1, Appli
	17	66	10.6	3294	3	US-08-685-467-1	Sequence 1, Appli
	18	66	10.6	3294	4	US-08-913-942-1	Sequence 1, Appli
	19	66	10.6	3294	4	US-09-268-347-43	Sequence 43, Appli
	20	66	10.6	5502	3	US-08-836-134-1	Sequence 1, Appli
	21	66	10.6	5502	4	US-09-493-784-1	Sequence 1, Appli
	22	65.5	10.5	7620	1	US-07-767-135-1	Sequence 1, Appli
	23	65.5	10.5	7620	1	US-07-841-652-1	Sequence 1, Appli
c	24	65	10.4	568	2	US-08-657-866-9	Sequence 9, Appli
c	25	65	10.4	568	4	US-09-163-951-9	Sequence 9, Appli
c	26	65	10.4	664	4	US-09-221-017B-469	Sequence 469, App
	27	65	10.4	5679	1	US-08-201-697-1	Sequence 1, Appli
	28	64.5	10.3	697	2	US-08-716-284-1	Sequence 1, Appli
c	29	64.5	10.3	1490	4	US-09-269-040-5	Sequence 5, Appli
c	30	64	10.2	321	4	US-09-240-274-103	Sequence 103, App
c	31	64	10.2	1000	1	US-08-599-252-96	Sequence 96, Appl
c	32	64	10.2	1000	5	PCT-US96-06352-96	Sequence 96, Appl
c	33	64	10.2	1000	5	PCT-US96-06583-96	Sequence 96, Appl
c	34	64	10.2	2025	4	US-09-247-155-58	Sequence 58, Appl
	35	64	10.2	2774	3	US-08-732-433-2	Sequence 2, Appli
	36	63.5	10.2	670	4	US-09-040-984-7	Sequence 7, Appli
	37	63.5	10.2	670	4	US-09-123-912-7	Sequence 7, Appli
c	38	63.5	10.2	670	4	US-09-643-597-7	Sequence 7, Appli
	39	63.5	10.2	670	4	US-09-221-017B-574	Sequence 574, App
	40	63.5	10.2	711	3	US-08-622-046B-2	Sequence 2, Appli
	41	63.5	10.2	711	3	US-08-622-046B-13	Sequence 13, Appl
	42	63.5	10.2	711	4	US-09-100-264-2	Sequence 2, Appli
	43	63.5	10.2	760	3	US-08-768-859A-7	Sequence 7, Appli
	44	63.5	10.2	760	3	US-08-767-820A-7	Sequence 7, Appli
	45	63.5	10.2	760	5	PCT-US95-06157-7	Sequence 7, Appli

```
QY 21 ArgGluLeuThrThrThrCysLeuAsnThrThrThrAlaSerValLeuGlnPheSerIle 40
|||||.....|
Db 1030 CGAGAGCTGACCACCATGCTGAACACAAACAGCATCTGCTCCAGTTTCCATT 1089

QY 41 GlySerGlySerCysArgPheSerTySerAspProSerIleThrValSerTyAlaLys 60
|||||.....|
Db 1090 GGGTCAGGATGATGCTGATTTAGTTACTCTGACCCAGCATCACTGTCTCATAGCCCAAG 1149

QY 61 AsnAsnThrAlaAspTrpIleGlnLeuGluLysIleArgAlaProSerAsnValSerThr 80
|||||.....|
Db 1150 AACAAATACCCCTGATGGATAGCTAGCGAGAAATAGAGCCCTTCCAAATGTGAGCACA 1209

QY 81 ValIleHisIleLeuTyLeuProGluGluAlaLysGlyGluSerValGlnPheGlnTrp 100
|||||.....|
Db 1210 GTCATCACATCTGTACCTCCCGAGGAAGCCAAAGGGAGAGCGTGCAGTTCCAGTGG 1269

QY 101 LysGlnAspSerLeuArgValGlyGluValTyGluAlaCysTrpAlaLeu 117
|||||.....|
Db 1270 AAACAGGACAGCTGGAGTGGGTGAGGTGTATGAGGCCCTGCTGGGCCCTG 1320

RESULT 2
US-09-334-220-4
; Sequence 4, Application US/09334220
; Patent No. 6323177
; GENERAL INFORMATION:
; APPLICANT: St. Judes Children's Research Hospital
; APPLICANT: Curran, Thomas
; APPLICANT: D'Arcangelo, Gabriella
; TITLE OF INVENTION: INTERACTION OF REELIN WITH VERY LOW
; TITLE OF INVENTION: DENSITY LIPOPROTEIN (VLDL) RECEPTOR FOR SCREENING AND
; TITLE OF INVENTION: THERAPIES
; FILE REFERENCE: 2427/0F704
; CURRENT FILING DATE: US/09/334,220
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 11580
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-334-220-4

Alignment Scores:
Pred. No.: 1,15e-68 Length: 11580
Score: 572.00 Matches: 106
Percent Similarity: 95.73% Conservative: 6
Best Local Similarity: 90.60% Mismatches: 5
Query Match: 91.52% Indels: 0
DB: 4 Gaps: 0

US-09-897-438B-2 (1-117) x US-09-334-220-4 (1-11580)
QY 1 GluGlnCysGlyThrThrIleMetHisGlyAsnAlaValThrPheCysGluProTyGlyPro 20
|||||.....|
Db 860 GAACAGTGTGGCCGATATATGATGCAATGCGCTCACCTCTGTGAACCATATGCCCCA 919

QY 21 ArgGluLeuThrThrThrCysLeuAsnThrThrThrAlaSerValLeuGlnPheSerIle 40
|||||.....|
Db 920 CGAAGACTGATTACCAAGCCCTTAATACAAACAGCTTCTCTCTCCAAATTTCCATT 979

QY 41 GlySerGlySerCysArgPheSerTySerAspProSerIleThrValSerTyAlaLys 60
|||||.....|
Db 980 GGGTCAGGTCATGTCGCTTTAGTTATTACAGCCCGAGCATCATCGTGTATATGCCAAG 1039

QY 61 AsnAsnThrAlaAspTrpIleGlnLeuGluLysIleArgAlaProSerAsnValSerThr 80
|||||.....|
Db 1040 AATAACTCTCGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1099

QY 81 ValIleHisIleLeuTyLeuProGluGluAlaLysGlyGluSerValGlnPheGlnTrp 100
|||||.....|
Db 1100 ATCATCCATATCTCTACCTTCTCTGAGGAGCCCAAGGGAGAGATGTCCTCAATTTTCAGTGG 1159

QY 101 LysGlnAspSerLeuArgValGlyGluValTyGluAlaCysTrpAlaLeu 117
|||||.....|
```

```
Db 1160 AACAGGAAATCTTCGTGTAGTGAAGTATATGAAGCCTGCTGGGCCTTA 1210

RESULT 3
US-09-351-814-12
; Sequence 12, Application US/09351814
; Patent No. 6352851
; GENERAL INFORMATION:
; APPLICANT: Nielsen, Bjarne Roenfeldt
; APPLICANT: Svendsen, Allan
; APPLICANT: Pedersen, Henrik
; APPLICANT: Vind, Jesper
; APPLICANT: Hendriksen, Hanne Vang
; APPLICANT: Frandsen, Torben Peter
; TITLE OF INVENTION: Glucoamylase Variants
; FILE REFERENCE: 5636.200-US
; CURRENT APPLICATION NUMBER: US/09/351,814
; CURRENT FILING DATE: 1999-07-12
; EARLIER APPLICATION NUMBER: PA 1998 00937
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: PA 1998 01667
; EARLIER FILING DATE: 1998-12-17
; EARLIER APPLICATION NUMBER: 60/093,528
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/115,545
; EARLIER FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 2602
; TYPE: DNA
; ORGANISM: ASPERGILLUS NIGER
US-09-351-814-12

Alignment Scores:
Pred. No.: 0.41 Length: 2602
Score: 81.00 Matches: 34
Percent Similarity: 36.30% Conservative: 15
Best Local Similarity: 25.19% Mismatches: 54
Query Match: 12.96% Indels: 32
DB: 4 Gaps: 4

US-09-897-438B-2 (1-117) x US-09-351-814-12 (1-2602)
QY 9 GlyAsnAlaValThrPheCysGluProTyGlyProArgGluLeuThrThrThrCysLeu 28
|||||.....|
Db 2007 GCGCGCACCATCTACGAGCGGTACCCCTGATCGGAGCGGTGACCTCGACCCAGCAAG 2066

QY 29 AsnThrThrThrAlaSerValLeuGlnPheSerIleGlySerGlySerCysArgPheSer 48
|||||.....|
Db 2067 ACCACCGGACTGCTAGCAAGACGACGACGACGACGACGACGACGACGACGACGACG 2126

QY 49 TyrSer-----AspProSerIleThrValSerTyAlaLysAsn----- 61
::: ||| ::| ::| ::| ::| ::| ::|
Db 2127 ACCGCGCTGGCTGTGACTTTTCGATCTGACAGCTTACCACCATCTACGCGGAGAACATCTAC 2186

QY 62 -----AsnThrAlaAspTrpIleGlnLeuGlu 70
|||||.....|
Db 2187 CTGCTCGGATCGATCTCTCAGCTGGGTGACTGGGAACACGCGCATAGCTCTGAGT 2246

QY 71 LysIleArgAlaProSerAsnValSerThrValIleHisIleLeuTyLeuProGluGlu 90
::: ||| ::| ::| ::| ::| ::| ::|
Db 2247 GCTGACAAAGTACACTTCTGACGAGCCCGCTCTGTGTATGTCTACTGTGACTCTGCGG----- 2300

QY 91 AlaLysGlyGluSerValGlnPhe-----Gln 99
|||||.....|
Db 2301 ---GCTGCTGAGTCGTTTGTAGTAAAGTTTATCCGCAATTGAGAGCGATGACTCCGTGGAG 2357

QY 100 TrpLysGlnAspSerLeuArgValGlyGluValTyGluAlaCys 114
|||||.....|
Db 2358 TGGGAGAGTGTATCCCAACCGAGAGATACACCGTTCCTCAGCGGTGC 2402

RESULT 4
```

US-08-905-223-151
; Sequence 151, Application US/08905223
; Patent No. 6222029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,223
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 151:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 414 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: CDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Brain
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 262..369
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 3.9
; OTHER INFORMATION: seq SICLSLLIPGPKP/LV
US-08-905-223-151

Alignment Scores:
Pred. No.: 0.0639 Length: 414
Score: 78.00 Matches: 28
Percent Similarity: 41.88% Conservative: 21
Best Local Similarity: 23.93% Mismatches: 48
Query Match: 12.48% Indels: 20
DB: 4 Gaps: 4

US-09-897-438B-2 (1-117) x US-08-905-223-151 (1-414)

QY 3 CysGlyThrIleMethHisGlyAsnAlaValThrPheCysGluProTyrGlyProArgGlu 22
Db 48 TGTGGCTCTACT-----GTTGATGCCCTGGTCTTATTGAAAGGCCACCCGGTAC 101
QY 23 LeuThrThrThrCysLeuAsnThrThrThrAlaSerValLeuGlnPheSerIleGlySer 42
Db 102 GTGTGCACACACAGCGTTCGCGTGAATTCCTTCTACAGATAGACTTC---GCT 158
QY 43 GlySerCysArgPheSerTyrSerAspProSerIleThrValSerTyrAlaLysAsnAsn 62
Db 159 GCCTCCGCTGCAGTCACAGACTCTGTGATGCGGATTAATGGAATACCTAGTAGATCTT 218
QY 63 ThrAlaAspTrpIleGlnLeuGluLysIleArgAlaProSerAsnValSer---ThrVal 81

Db 219 GGATTGTGCATGGCACCACCTGGTAAAGGACTGTCTGCTACCAATGTGAATGCAGTGC 278
QY 82 IleHisIle-----Leu 85
Db 279 TATCATCTGCAACGGATCCTGGTGTCTAGACACDWTCAACAGTGGACTAGAACACTCTG 338
QY 86 TyrLeuProGluGluAlaLysGlyGluSerValGlnPheGlnTrpLysGln 102
Db 339 CCTCTCCTCTCTTATACCAGGTCCCAAGCCACTCGTTTCGTTGGGCATCAA 389

RESULT 5

US-09-336-643A-17
; Sequence 17, Application US/09336643A
; Patent No. 6396761
; GENERAL INFORMATION:
; APPLICANT: Miller, Andrew P.
; APPLICANT: Curran, Mark Edward
; APPLICANT: Hu, Ping
; APPLICANT: Rutter, Marc
; APPLICANT: Wang, Jian-Wang
; TITLE OF INVENTION: No. 6399761el Human Potassium Channels
; FILE REFERENCE: SEQ-15P
; CURRENT APPLICATION NUMBER: US/09/336,643A
; CURRENT FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 60/076,687
; PRIOR FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: 60/116,448
; PRIOR FILING DATE: 1999-01-19
; PRIOR APPLICATION NUMBER: PCT/US99/03826
; PRIOR FILING DATE: 1999-02-22
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 3102
; TYPE: DNA
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (274)...(1705)
; OTHER INFORMATION: K+Hnov11
US-09-336-643A-17

Alignment Scores:
Pred. No.: 8.26 Length: 3102
Score: 72.50 Matches: 34
Percent Similarity: 37.01% Conservative: 13
Best Local Similarity: 26.77% Mismatches: 56
Query Match: 11.60% Indels: 25
DB: 4 Gaps: 5

US-09-897-438B-2 (1-117) x US-09-336-643A-17 (1-3102)

QY 2 GlnCysGlyThrIleMethHisGlyAsnAlaValThrPheCysGluProTyrGlyProArg 21
Db 336 CAATGTGGCGGCTTCAAGAGAGGCTGCGCTCCACACACGCTGCTGGCT-TCCCCGAGA 394
QY 22 GluLeuThrThrThrCysLeuAsnThrThrAla-----SerValLeu 36
Db 395 CGCGCCTGGCGCGCTTGTCTCTGCCACTCGCGAGGCCATTTCTGGAGCTTGGCATG 454
QY 37 GlnPheSerIleGlySerGlySerCysArgPheSer-----TyrSerAspProSer 53
Db 455 ACTACGACGACGCTCCACGCGGAGTTCCTACGACCCGACCCGAGAGCTTCCCCCTACG 514
QY 54 IleThrValSerTyrAlaLysAsnAsnThrAlaAspTrpIleGlnLeuGluLysIleArg 73
Db 515 TGCTGCATTCTATCACACCGGCAAGCTTCACGTCATGG-----CTGAGCTATGTCTCT 568
QY 74 AlaProSer-----AsnValSerThr-ValIleHisIleLeuTyrLe 87
Db 569 TCCTCTTCAGCCAGGAGATCGAGTACTTGGGGCATCAACAGGATCTTCTCATTTGACTCCTGCT 628


```

; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 265:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2231 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-527-265

Alignment Scores:
Pred. No.: 21.2 Length: 2231
Score: 68.00 Matches: 30
Percent Similarity: 37.70% Conservative: 16
Best Local Similarity: 24.59% Mismatches: 42
Query Match: 10.88% Indels: 34
DB: Gaps: 6

US-09-897-438B-2 (1-117) x US-08-961-527-265 (1-2231)
Qy 6 IleMethisGlyAsnAlaValThrPheCysGluPro----- 17
Db 577 ATCTTTAGTAAGACAGAGCTCTGTTTCACGAGCCCTATTTCCAAGTTACATTTGCGCCAT 518
Qy 18 -----TyrGlyProArgGluLeuThrThrCysLeuAsnThrThrAla--- 33
Db 517 CTAATTAGCTATGCTAAAAAAGAGCGCAGGATGCTCTAGGACCAAGGATGCCATG 458
Qy 34 -----SerValLeuGlnPheSerIleClySerGlySerCysArgPheSerTyrSer 50
Db 457 TTGGTTCTAAATCATGCTCTTTGGTCGGTCTCTTTTCCCAACGAATAGTCGCTTC 398
Qy 51 AspProSerIle-----ThrValSerTyrAlaLysAsnThrAlaAspTrpIle 67
Db 397 GTTCCCTCTGTTTAACTCGCGAGGTGAGTCAGTCTTTTAATCGC----- 353
Qy 68 GlnLeuGluLysIleArgAlaProSerAsnValSerThrValIleHisIle----- 84
Db 352 ATGTCACGACAGGTT---GTTCCCAAAAGAGAGACGACCTGCTTACCTGATGAATCAG 296
Qy 85 -----LeuTyrLeuProGluAlaLysGlyGluSerValGln 97
Db 295 CCTATCTACCAAGTTTGTATGCTCATGCACACAGAGAATCTCGAAGCGCGCAGCTGAT 236
Qy 98 PheGln 99
Db 235 TTTGAA 230

```

RESULT 10
5208144-7
; Patent No. 5208144
; APPLICANT: SMITH, JOHN A.; RAYCHORDHURY, RAKTIMA; NILES, JOHN L.
; TITLE OF INVENTION: METHOD FOR DETECTION OF HUMAN DNA
; CONTAINING THE GENE ENCODING LOW DENSITY LIPOPROTEIN RECEPTOR
; NUMBER OF SEQUENCES: 42
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/396,697
; FILING DATE: 22-AUG-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 313,682
; FILING DATE: 22-FEB-1989
; APPLICATION NUMBER: 235,211
; FILING DATE: 23-AUG-1988
; SEQ ID NO: 7:
; LENGTH: 2937
5208144-7

Alignment Scores:
Pred. No.: 37.9 Length: 2937
Score: 67.50 Matches: 28
Percent Similarity: 36.50% Conservative: 22
Best Local Similarity: 20.44% Mismatches: 59
Query Match: 10.80% Indels: 28
Gaps: 6
US-09-897-438B-2 (1-117) x 5208144-7 (1-2937)
Qy 1 GluGlnCysGlyThrIleMethHisGlyAsnAlaValThrPheCysGluPro---TyrGly 19
Db 682 CAGACTGTACCACGCTGACGAGCGGCGTTCATCTCTCTCGACACCGGGTTCAA 741
Qy 20 ProArgGluLeu---ThrThrThrCysLeuAsnThrThrAlaSerValLeuGlnPhe 38
Db 742 CCCAGTACTTTGGACAAAATTCCTGTCAAGACATCAATGAATGTGAG-----GAGTTT 795
Qy 39 SerIleGlySerGlySerCysArgPheSerTyrSerAsp----- 51
Db 796 GCATCTGCCCGCAGAGTGTGGAACAGCAAGAAAGTATGATGTTCTGTGTCTCAT 855
Qy 52 -----ProSerIleThrValSerTyrAlaLysAsnAsnThrAlaAspTrp----- 66
Db 856 GCGTTCAAGTCTATCAGTACCATATATGAGAACGGTGTGAGCTGTGGAAGCCCTCT 915
Qy 67 -----IleGlnLeuGluLysIleArgAlaProSerAsnValSer 79
Db 916 CTCCTGCTCCTCGCGAAGTGTCCGAATCCGGAAGTACATACCTCTCTGAGAATGTC 975
Qy 80 ThrValIleHisIleLeuTyrLeuProGluGluAlaLysGlyGluSerValGlnPheGln 99
Db 976 TCAGAG-----TACTGGAAGAGGAGGAGCATATCCAACTATTGACTATGAC 1023
Qy 100 TrpLysGlnAspSerLeuArgValGlyGluValTyr-GluAlaCysTrp 115
Db 1024 TGGATCCCGACACATAGGCTCAGTGTGTCTATTACACTGTCTCG 1072

RESULT 11
US-09-134-001C-2455/c
; Sequence 2455, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2455
; LENGTH: 741
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2455
Alignment Scores:
Pred. No.: 5.37 Length: 741
Score: 67.00 Matches: 20
Percent Similarity: 47.37% Conservative: 7
Best Local Similarity: 35.09% Mismatches: 16
Query Match: 10.72% Indels: 14
Gaps: 3
US-09-897-438B-2 (1-117) x US-09-134-001C-2455 (1-741)
Qy 9 GlyAsnAlaValThrPheCysGluProTyrGlyProArgGluLeuThrThrCysLeu 28
Db 380 GGAGATGCTTTTCTTATTGCTAAACCGATAGACCATATCCA----- 339
Qy 29 AsnThrThrThrAlaSerValLeuGlnPheSerIleGlySerGlySerCysArgPhe--- 47
Db 338 CATCTACATCCGCGATGGTTTA-----CTTGGCGCTGGAGGATGTTCTTTCAA 288
Qy 48 -----SerTyrSerAspProSerIleThrValSerTyrAlaLys 60
Db 287 AAGTGTTAATAGTAGTCTGATCCAAATCCACATATCTTTGAAAAA 237
RESULT 12
US-08-104-158-1/c
; Sequence 1, Application US/08104158
; Patent No. 6215042
; GENERAL INFORMATION:
; APPLICANT: Willmitzer, Lothar
; APPLICANT: Sonnewald, Uwe
; APPLICANT: Kossmann, Jens
; APPLICANT: Mueller-Roeber, Bernd
; APPLICANT: Visser, Richard Gerardus Franciscus
; APPLICANT: Jacobsen, Evert
; TITLE OF INVENTION: PLASMIDS CONTAINING DNA-SEQUENCES THAT
; TITLE OF INVENTION: CAUSE CHANGES IN THE CARBOHYDRATE CONCENTRATION AND THE
; TITLE OF INVENTION: CARBOHYDRATE COMPOSITION IN PLANTS, AS WELL AS PLANT CELLS
; TITLE OF INVENTION: AND PLANTS CONTAINING THESE PLASMIDS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen
; STREET: 1180 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-8403
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/104,158
; FILING DATE: 13-AUG-1993
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP PCT/EP92/00302
; FILING DATE: 11-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 41 04 782.6
; FILING DATE: 13-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meilman, Edward A.
; REGISTRATION NUMBER: 24,735
; REFERENCE/DOCKET NUMBER: FA-1996 PCT (951-91)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-382-0700


```
; TELEFAX: 212-382-0888
; TELEX: 236925
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2909 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Solanum tuberosum cv. Desiree
; STRAIN: Desiree
; DEVELOPMENTAL STAGE: growing tuber
; TISSUE TYPE: tuber
; CELL TYPE: total tuber
; IMMEDIATE SOURCE:
; LIBRARY: cDNA of total tuber mRNA in pUC 19 (Hinc II)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1699
; OTHER INFORMATION: /note= "for Branching enzyme I
; OTHER INFORMATION: (partial) truncated protein; 97,11 % identity to
; OTHER INFORMATION: active potato branching enzyme"
;
US-08-104-158-1

Alignment Scores:
Pred. No.: 51.5 Length: 2909
Score: 66.50 Matches: 28
Percent Similarity: 45.79% Conservatve: 21
Best Local Similarity: 26.17% Mismatches: 37
Query Match: 10.64% Indels: 21
DB: 4 Gaps: 6

US-09-897-438b-2 (1-117) x US-08-104-158-1 (1-2909)
Qy 26 ThrCysLeuAsnThrThrAlaSerVal-----LeuGlnPheSer----- 39
Db 1589 ACATGCCAGATACATCTCTTTGTCATAGGAGAAATGCAATGCTGTGCACCGACAA 1530
Qy 40 ---lleGlySerGlySerCysArgPheSerTyrSerAspProSerIleThrValSerTyr 58
Db 1529 TAGACTGATCGCTCTCCGCATATGCTATACACT-----TCTCTGATATCTCTAT 1476
Qy 59 AlAlAsnAsn-----ThrAlaAspTrpIleGlnLeuGluLysIleArgAlaProSer 76
Db 1475 TTGTCAAACTCGATGTTACTTCTTCATGGACCAATCTTCATCATCTTATTTCTTAAAT 1416
Qy 77 AsnValSerThrValIleHisIle-----LeuTyrLeuPro 88
Db 1415 AATCTATCCACTATCTGCGGATGCCATGTCAGCGGATGATCAAAACCAATTCCTCCT 1356
Qy 89 GluGluAlaLys-----GlyGluSerValGlnPheGlnTrpLysGlnAspSerLeuArg 106
Db 1355 CAGAAACAGGCGCGCTAAGGCGCGCATACCAAGAAACATCTTCGCGCAATACACTTTCGT 1296
Qy 107 ValGlyGluValTyrGluAla 113
Db 1295 CTGGGAAATCTTGTGAATCA 1275

RESULT 13
US-09-240-274-214/c
; Sequence 214, Application US/09240274
; Patent No. 6253455
; GENERAL INFORMATION:
; APPLICANT: Siegel, Donald L.
; TITLE OF INVENTION: Rh(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL
; TITLE OF INVENTION: SORTING METHOD FOR PRODUCTION THEREOF
; FILE REFERENCE: 09596-4202
; CURRENT APPLICATION NUMBER: US/09/240, 274
; CURRENT FILING DATE: 1999-01-29
; EARLIER APPLICATION NUMBER: 60/081,380
```

```
; EARLIER FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/028,550
; EARLIER FILING DATE: 1996-10-11
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 214
; LENGTH: 318
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: anti-Rh(D) antibody clone SH44
;
US-09-240-274-214

Alignment Scores:
Pred. No.: 2.02 Length: 318
Score: 66.00 Matches: 25
Percent Similarity: 38.32% Conservatve: 16
Best Local Similarity: 23.36% Mismatches: 48
Query Match: 10.56% Indels: 18
DB: 4 Gaps: 4

US-09-897-438b-2 (1-117) x US-09-240-274-214 (1-318)
Qy 15 CysGluProTyrGlyProArgGluLeuThrThrThrCysLeuAsnThrThrThrAlaSer 34
Db 303 TGTGTCCTCCAGGCGCGAAAGAGAGATTACTTAACCTCTGTGGCAGTAGTAAGTTGGAA 244
Qy 35 ValLeuGlnPheSerIleGlySerGlySer-----CysArgPheSerTyrSer--- 50
Db 243 ATCTTCAGGTTCGACACTGCTGTGAGAGTGAACCTGTCTCCAGATCCACTGGCACT 184
Qy 51 -----AspProSerIleThrValSerTyrAla-----LysAsnAsnThr 63
Db 183 GAACCTTGATGGGACCCCACTTTGTAACCTGGATATATCTATCATCAGGAGACAGGGC 124
Qy 64 AlaAspTrpIleGlnLeuGluLysIleArgAlaProSerAsnValSerThrValIleHis 83
Db 123 TTTTCCAGGCTCTGTTGATACCAATTCAAGAACCTGGGAATGCTCTGACTTCG----- 70
Qy 84 IleLeuTyrLeuProGluGluAlaLysGlyGluSerValGlnPheGlnTrpLysGlnAsp 103
Db 69 -----CCGCAAGTGATGATGACTCTGTCTCTACAGATCGACAGGGA 25
Qy 104 SerLeuArgValGlyGluVal 110
Db 24 CGATGGAGACTGGGTGAGCTC 4

RESULT 14
US-08-791-347-2
; Sequence 2, Application US/08791347
; Patent No. 5885776
; GENERAL INFORMATION:
; APPLICANT: Stone, Edwin M.
; APPLICANT: Sheffield, Val C.
; APPLICANT: Alward, Wallace L.M.
; TITLE OF INVENTION: GLAUCOMA COMPOSITIONS AND THERAPEUTIC
; TITLE OF INVENTION: AND DIAGNOSTIC USES THEREFOR
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESS: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/791,347
; FILING DATE: 30-JAN-1997
```

```
;
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: UIA-010.26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1509 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-791-347-2

Alignment Scores:
Pred. No.: 22.1 Length: 1509
Score: 66.00 Matches: 24
Percent Similarity: 44.86% Conservative: 24
Best Local Similarity: 22.43% Mismatches: 35
Query Match: 10.56% Indels: 24
Gaps: 5

US-09-897-438B-2 (1-117) x US-08-791-347-2 (1-1509)
Qy 15 CysGluProTyrGlyPro-----ArgGluLeuThrThrCysLeu--- 28
Db 22 TGTCCAGCTTTGGGCGCTGAGATGCCAGCTGCCAGCTGCTTCCGCTGGCTGGTG 81
Qy 29 -----AsnThrThrAlaSerValLeuGlnPheSerIleGlySerGlySerCys 45
Db 82 TGGGATGTGGGGCGGAGACAGCTCAGCTCAGGAAGGCAATGACAGAGTGGCGGATGC 141
Qy 46 ArgPheSerTyrSer-----AspPro----- 52
Db 142 CAGTATACCTTCAGTGTGGCGGAGTCCCAATGAATCCAGCTGCCAGAGCAGAGGCC 201
Qy 53 ---SerIleThrValSerTyrAlaIysAsnAsnThrAlaAspTrpIleGlnLeuGluLys 71
Db 202 ATGCAGTCATCCATACTTACAGAAAGACAGACAGACCCCAACGCTTAGAGCTGGAGGCC 261
Qy 72 IleArgAlaPro---SerAsnValSerThrValIleHisIleLeuTyrLeuProGluGlu 90
Db 262 ACCAAAGCTCGACTCAGCTCCCTGGAGAGCGCTCCTCCACCAATGACCTTGGACCGGCT 321
Qy 91 AlalysGlyGluSerValGln 97
Db 322 GCCAGGCCCGGAGAGAGCCAG 342

RESULT 15
US-08-791-347-1
; Sequence 1, Application US/08791347
; Patent No. 5885776
; GENERAL INFORMATION:
; APPLICANT: Stone, Edwin M.
; APPLICANT: Sheffield, Val C.
; APPLICANT: Alward, Wallace L.M.
; TITLE OF INVENTION: GLAUCOMA COMPOSITIONS AND THERAPEUTIC
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
;
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/791,347
; FILING DATE: 30-JAN-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: UIA-010.26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1985 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-791-347-1

Alignment Scores:
Pred. No.: 33.6 Length: 1985
Score: 66.00 Matches: 24
Percent Similarity: 44.86% Conservative: 24
Best Local Similarity: 22.43% Mismatches: 35
Query Match: 10.56% Indels: 24
Gaps: 5

US-09-897-438B-2 (1-117) x US-08-791-347-1 (1-1985)
Qy 15 CysGluProTyrGlyPro-----ArgGluLeuThrThrCysLeu--- 28
Db 89 TGTCCAGCTTTGGGCGCTGAGATGCCAGCTGCCAGCTGCTTCCGCTGGCTGGTG 148
Qy 29 -----AsnThrThrAlaSerValLeuGlnPheSerIleGlySerGlySerCys 45
Db 149 TGGGATGTGGGGCGGAGACAGCTCAGCTCAGGAAGGCAATGACAGAGTGGCGGATGC 208
Qy 46 ArgPheSerTyrSer-----AspPro----- 52
Db 209 CAGTATACCTTCAGTGTGGCGGAGTCCCAATGAATCCAGCTGCCAGAGCAGAGGCC 268
Qy 53 ---SerIleThrValSerTyrAlaIysAsnAsnThrAlaAspTrpIleGlnLeuGluLys 71
Db 269 ATGTCATCATCCATACTTACAGAAAGACAGACAGACCCCAACGCTTAGACCTGGAGGCC 328
Qy 72 IleArgAlaPro---SerAsnValSerThrValIleHisIleLeuTyrLeuProGluGlu 90
Db 329 ACCAAAGCTCGACTCAGCTCCCTGGAGAGCGCTCCTCCACCAATGACCTTGGACCGGCT 388
Qy 91 AlalysGlyGluSerValGln 97
Db 389 GCCAGGCCCGGAGAGAGCCAG 409

Search completed: November 6, 2002, 20:54:57
Job time : 76 secs
```

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 6, 2002, 18:12:02 : Search time 299 Seconds
(without alignments)
881.217 Million cell updates/sec

Title: US-09-897-438B-2
Perfect score: 625
Sequence: 1 EQCGTIMGNAVTCEPYGP.....FQWKQDSLRVGEVYACWAL 117

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delopt 6.0, Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DRV=xlp
-Q=/cgn2_1/USPTO.spool/US09897438/runat_06112002_101923_11281/app_query.fasta_1.263
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPTXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09897438 @CGL1.1.79 runat_06112002_101923_11281 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	625	100.0	351	24	ABL40165	Mouse reelin prote
2	625	100.0	2745	24	ABD22754	Mus musculus trunc
3	625	100.0	11673	24	ABA92603	Mouse reelin encod
4	625	100.0	11673	24	ABN99284	Mouse ischaemic co
5	572	91.5	11580	24	ABN96939	Gene #3437 used to
6	572	91.5	11580	24	ABA92604	Human reelin encod
7	572	91.5	11632	23	AAS89484	DNA encoding novel
8	499	79.8	2025	24	AAD22779	Xenopus laevis int
9	499	79.8	2274	24	AAD22753	Xenopus laevis int
c 10	144	23.0	413	22	AAK07720	Human brain expres
c 11	144	23.0	413	22	AAK33553	Human bone marrow
c 12	144	23.0	413	24	ABS08422	Human genome-deriv
c 13	109.5	17.5	4576	23	AAS89475	DNA encoding novel
c 14	107	17.1	87	22	AAK20384	Human brain expres
c 15	107	17.1	87	22	AAK46490	Human bone marrow
c 16	107	17.1	87	24	ABS20836	Human genome-deriv
c 17	83	13.3	1174	21	AAF11251	Aspergillus niger
c 18	83	13.3	1845	5	AAAN40165	Sequence encoding
c 19	83	13.3	3408	5	AAAN40166	Sequence of A.awam
c 20	81.5	13.0	1404	23	AAS89482	DNA encoding novel
c 21	81	13.0	2258	21	AAF11248	Aspergillus niger
c 22	81	13.0	2600	9	AAAN2019	Sequence of amylog
c 23	81	13.0	2602	21	AAZ87843	A. niger G1 glucosa
c 24	81	13.0	2602	22	AAC85109	A. niger nucleotid
c 25	80	12.8	9477	23	ABL14686	Drosophila melanog
c 26	78	12.5	414	20	AAK51900	Human secreted pro
c 27	73.5	11.8	700	22	AAH92336	Human inflammatory
c 28	73	11.7	1820	11	AAQ04731	CDNA sequence from
c 29	73	11.7	2615	11	AAQ04730	Sequence encoding
c 30	72.5	11.6	3102	20	AAZ11905	Human potassium ch
c 31	72.5	11.6	28473	19	AAV52216	Streptococcus pneu
c 32	72	11.5	795	24	ABK75454	Bacillus lichenifo
c 33	72	11.5	1654	21	AAK56613	ECORI/XbaI fragmen
c 34	71.5	11.4	932	23	ABV22646	Human prostate exp
c 35	71.5	11.4	932	23	ABV28470	Human prostate exp
c 36	71.5	11.4	2153	23	AAK73449	DNA encoding novel
c 37	71	11.4	586	21	AAF11295	Aspergillus niger
c 38	71	11.4	2562	22	AAC84443	Thermosacus crusta
c 39	70.5	11.3	813	24	ABN70601	Streptococcus poly
c 40	70.5	11.3	837	24	ABN66616	Streptococcus poly
c 41	70.5	11.3	871	23	ABV23361	Human prostate exp
c 42	70.5	11.3	871	23	ABV29216	Human prostate exp
c 43	70.5	11.3	1237	21	AAC51174	Arabidopsis thalia
c 44	70.5	11.3	1239	21	AAC47796	Arabidopsis thalia
c 45	70	11.2	531	22	AAI13506	Human breast cance

ALIGNMENTS

RESULT 1
ABL40165
ID ABL40165 standard; DNA; 351 BP.
XX
AC ABL40165;
XX
DT 21-MAY-2002 (first entry)
XX
DE Mouse reelin protein CR-50 epitope region encoding DNA SEQ ID NO:1.
XX
KW Mouse; reelin protein CR-50 epitope region; elucidation; neuron;
XX cerebral disturbance; reelin protein; neuroprotective; gene; ds.
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 1..351
FT /*tag= a

```
FT /partial
FT /product= "reelin protein CR-50 epitope region"
FT /note= "no start or stop codons given"
XX
XX JP2002017361-A.
XX
XX 22-JAN-2002.
XX
XX 04-JUL-2000; 2000JP-0202801.
XX
XX 04-JUL-2000; 2000JP-0202801.
XX
XX (RIKA ) RIKAGAKU KENKYUSHO.
XX
XX WPI; 2002-221707/28.
XX
XX P-PSDB; ABB06244.
XX
XX Reelin protein CR-50 epitope region, useful for diagnosis and treatment
XX of cerebral disturbance
XX
XX Claim 5; Page 11; 16pp; Japanese.
XX
XX The present sequence encodes the mouse reelin protein CR-50 epitope
XX region, which contains the CR-50 antibody recognition site and is free
XX from F-spondin domains and repetitive sites. Also described are: (1) an
XX expression vector comprising a polynucleotide encoding a reelin protein
XX epitope region; (2) host cells with transfected the expression vector;
XX (3) polypeptides prepared by culture of the host cells; and (4)
XX polynucleotides comprising the 351 base sequence given in ABL40165 which
XX encodes the 117 amino acid sequence given in ABB06244; and (5) use of
XX the polynucleotide for diagnosis and/or treatment of diseases caused by
XX abnormal positioning of neural cells, and stimulation of association of
XX reelin protein. The mouse reelin protein CR-50 epitope region has
XX neuroprotective activity, and can be used in the diagnosis and treatment
XX of cerebral disturbance due to an abnormal reelin gene and positioning
XX of neurons.
XX
XX Sequence 351 BP; 86 A; 98 C; 92 G; 75 T; 0 other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 9.49e-72 Length: 351
XX Score: 625.00 Matches: 117
XX Percent Similarity: 100.00% Conservativeness: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 24 Gaps: 0
XX
XX US-09-897-438B-2 (1-117) x ABL40165 (1-351)
XX
Qy 1 GluGlnCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20
Db 1 GAGCAGTGTGGCACCATCATCGTGGCAATGCTGCACCTTCTGTGAGCCGTAGCGCCT 60
Qy 21 ArgGluLeuThrThrThrCysLeuAsnThrThrThrAlaSerValLeuGlnPheSerIle 40
Db 61 CGAGAGTGACACACCATGCTGACACACACACAGCATCTGCCCTCCAGTTTCCATT 120
Qy 41 GlySerGlySerCysArgPheSerThrSerAspProSerIleThrValSerTyrAlaLys 60
Db 121 GGGTCAGGATCATCTCGATTAGTTACTCTGACCCAGCATCACTGTCTATAGCCCAAG 180
Qy 61 AsnAsnThrAlaAspTrpIleGlnLeuGluLysIleArgAlaProSerAsnValSerThr 80
Db 181 AACAAATACCGCTGATTGGATTGAGTTCAGTGGAGAAATTAGAGCCCTTCCAATGTGAGCACA 240
Qy 81 ValIleHisIleLeuThrLeuProGluGluAlaLysGlyGluSerValGlnPheGlnTrp 100
Db 241 GTCATCACATCTCTGTACCTCCCGAGAGAGCCAAAGGGAGAGCGTGCAGTTCACAGTGG 300
Qy 101 LysGlnAspSerLeuArgValGlyGluValTyrGluAlaCysTrpAlaLeu 117
Db 301 AACACAGGACAGCCTGCGAGTGGGTGAGGTGTATGAGGCCCTGCTGGGCCCTG 351
```

```
RESULT 2
AAD22754
ID AAD22754 standard; cDNA; 2745 BP.
XX
XX AAD22754;
XX
XX 26-FEB-2002 (first entry)
XX
XX Mus musculus truncated reelin cDNA.
XX
XX Mouse; reelin; F-spondin domain; CR-50 epitope; gene therapy; agyria;
XX polymicrogyria; ectopic gray matter; ss.
XX
XX Mus musculus.
XX
XX Key Location/Qualifiers
XX CDS 283..2052
XX /*tag= a
XX /product= "Mouse truncated reelin protein"
XX sig_peptide 283..363
XX /*tag= b
XX mat_peptide 364..2049
XX /*tag= c
XX /product= "Mature truncated reelin protein"
XX misc_feature 284..849
XX /*tag= d
XX /note= "Encodes F-spondin domain"
XX misc_feature 970..1320
XX /*tag= e
XX /note= "Encodes CR-50 epitope region"
XX
XX EPI149844-A2.
XX
XX 31-OCT-2001.
XX
XX 11-APR-2001; 2001EP-0303411.
XX
XX 11-APR-2000; 2000JP-0109954.
XX
XX (RIKE ) RIKEN KK.
XX
XX Mikoshiba K, Tabata H, Nakajima K;
XX
XX WPI; 2002-019320/03.
XX
XX P-PSDB; AAE13606.
XX
XX Novel truncated Reelin protein containing F-spondin domain and CR-50
XX recognition site of Reelin protein, but not having Reelin repeat site,
XX useful to treat diseases including agyria due to abnormal neuron
XX alignment
XX
XX Claim 10; Page 20-26; 47pp; English.
XX
XX The invention relates to a truncated Reelin protein comprising a
XX F-spondin domain and a CR-50 recognition site but no reelin protein
XX repeat site. Reelin is an essential molecule in developing a normal
XX laminated structure of cerebrum. The truncated reelin protein and its
XX DNA are useful for treating diseases including agyria, polymicrogyria,
XX and ectopic gray matter due to abnormal neuronal alignment. Truncated
XX reelin protein DNA is useful in gene therapy. The present sequence is
XX a cDNA encoding Mus musculus truncated reelin protein.
XX
XX Sequence 2745 BP; 661 A; 716 C; 714 G; 654 T; 0 other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 1.75e-70 Length: 2745
XX Score: 625.00 Matches: 117
XX Percent Similarity: 100.00% Conservativeness: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 24 Gaps: 0
XX
XX US-09-897-438B-2 (1-117) x AAD22754 (1-2745)
```


PF 18-MAY-2001; 2001WO-JP04192.
XX
PR 18-MAY-2000; 2000JP-0145977.
XX
PA (UYN1-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX
XX WPI; 2002-034733/04.
DR - P-PSDB; ABB57065.
XX
PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or
PT by determining the expression profile of a gene group comprising these
PT genes -
XX
PS Claim 2; Page 362-385; 2690pp; English.
XX
CC The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC expression levels of particular genes (ABI99202 to ABI99912, encoding
CC the protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The
CC expression levels or expression profiles produced by these genes are
CC used as an indicator when screening for ischaemic condition-improving
CC drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914
CC represent PCR primers for a mouse ischaemic condition related sequence,
CC which are used in the exemplification of the present invention.
XX
SQ Sequence 11673 BP; 2831 A; 2985 C; 2985 G; 2872 T; 0 other;

Alignment Scores:
Pred. No.: 1.37e-69 Length: 11673
Score: 625.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

US-09-897-438B-2 (1-117) x ABI99284 (1-11673)

Qy 1 GluGlnCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20
Db 970 GAGCAGTGTGCACCATCATGCGCAATGCTGTCTCACCCTTCTGTGAGCCGTACGGCCCT 1029

Qy 21 ArgGluLeuThrThrCysLeuAsnThrThrThrAlaSerValLeuGlnPheSerile 40
Db 1030 CGAGAGCTGACCACCATGCTGAACACAAACAGCATCTGCTCCAGTTTCCATT 1089

Qy 41 GlySerGlySerCysArgPheSerTyrSerAspProSerIleThrValSerTyrAlaLys 60
Db 1090 GGGTCAGGATCATGCTGATTTAGTTACTGTGACCCAGCATCACTGTGTATAGCCCAAG 1149

Qy 61 AsnAsnThrAlaAspTptIleGlnLeuGluLysIleArgAlaProSerAsnValSerThr 80
Db 1150 AACAAATACCGCTGATTTGGATTCAGCTCGAGAAATTAAGACCCCTTCCAATGTGAGCACA 1209

Qy 81 ValIleHisIleLeuTyrLeuProGluGluAlaLysGlyGluSerValGlnPheGlnTrp 100
Db 1210 GTCATCCACATCCTGTACCTCCCGAGGAGGCCAAAGGGGAGACGCTGTCAGTTCCAGTGG 1269

Qy 101 LysGlnAspSerLeuArgValGlyGluValTyrGluAlaCysTrpAlaLeu 117
Db 1270 AAACAGGACAGCCTGCGAGTGGGTGAGGTGTATGAGGCCTGCTGGGCCCTG 1320

RESULT 5
ABN96939
ID ABN96939 standard; DNA; 11580 BP.
XX

AC ABN96939;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #3437 used to diagnose liver cancer.
XX
KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumour; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
PN WO200229103-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US30589.
XX
PR 02-OCT-2000; 2000US-237054P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX
DR WPI; 2002-426119/45.
XX
PT Diagnosing and detecting the progression of liver cancer,
PT hepatocellular carcinoma or metastatic liver tumor in a patient,
PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample -
XX
PS Claim 1; SEQ ID NO 3437; 298pp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 11580 BP; 3014 A; 2696 C; 2753 G; 3116 T; 1 other;

Alignment Scores:
Pred. No.: 1e-62 Length: 11580
Score: 572.00 Matches: 106
Percent Similarity: 95.73% Conservative: 6
Best Local Similarity: 90.60% Mismatches: 5
Query Match: 91.52% Indels: 0
DB: 24 Gaps: 0

US-09-897-438B-2 (1-117) x ABN96939 (1-11580)

Qy 1 GluGlnCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20
Db 860 GAACAGTGTGGCGGATTTATGATGCGAATGCGGTCACTTCTGTGAACCATATGGCCCA 919

Qy 21 ArgGluLeuThrThrCysLeuAsnThrThrThrAlaSerValLeuGlnPheSerile 40
Db 920 CGAGAACTGATTACACAGGCTTAATACAAACAGCTTCTGTCTCCCAATTTCCATT 979

Qy 41 GlySerClySerCysArgPheSerTyrSerAspProSerIleThrValSerTyrAlaLys 60
Db 980 GGGTCAGGTTTCATGTCGCTTAGTTATTATGAGACCCCAAGCATCATCGTGTATATGCAAG 1039

QY 61 AsnAsnThrAlaAspTrpIleGlnLeuGluLysIleArgAlaProSerAsnValSerThr 80
 Db 1040 AATAACTCTGCGACTGGATTACAGTCAGTAGAGAAAATTTAGAGCCCTTCCCAATGTCAGCACCA 1099
 QY 81 ValIleHisIleLeuTyrLeuProGluGluAlaLysGlyGluSerValGlnPheGlnTrp 100
 Db 1100 ATCATCCATATCTCTACCTTCTGAGGACGCCAAAGGGGAGAAATGTCCAATTTTCAGTGG 1159
 QY 101 LysGlnAspSerLeuArgValGlyGluValTyrGluAlaCysTrpAlaLeu 117
 Db 1160 AAGCAGGAAAATCTTCTGTAGGTGAAGTGTATGAAGCCCTGTCGGCCCTTA 1210
 RESULT 6
 ID ABA92604 standard; cdna; 11580 BP.
 AC ABA92604;
 XX 21-MAR-2002 (first entry)
 DE Human reelin encoding cDNA SEQ ID NO:4.
 KW Human; reelin; low density lipoprotein receptor; LDLR; neuroprotective;
 KW extracellular glycoprotein; neotropic; antilipaeamic; Alzheimer's disease;
 KW neurodegenerative disorder; neuronal regeneration; cognitive function;
 KW lipid metabolism disease; memory; developmental disorder; gene; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 CDS 176..10558
 FT /*tag= a
 FT /*product= "human reelin"
 XX
 PN US623177-B1.
 PD 27-NOV-2001.
 PF 16-JUN-1999; 99US-0334220.
 XX 16-JUN-1999; 99US-0334220.
 PR (SUSD-) ST JUDE CHILDREN'S RES HOSPITAL.
 PA Curran T, D'Arcangelo G;
 PI WPI; 2002-096596/13.
 DR P-PSDB; ABB05007.
 XX
 PT Novel composition useful for screening compounds that modulate Reelin
 PT binding to low density lipoprotein receptor, comprising an isolated
 PT Reelin polypeptide and low density lipoprotein receptor -
 XX
 PS Example 1; Column 75-84; 45pp; English.
 XX
 CC The present invention describes a composition (I) comprising an
 CC isolated reelin protein (II) bound to an isolated low density lipoprotein
 CC receptor (LDLR) (III). (II) is an extracellular glycoprotein of
 CC approximately 385 kda containing a small region of similarity with
 CC F-spondin at the N terminus, a stretch of positively charged amino
 CC acids at the C terminus, and a series of eight internal repeats of
 CC 350-390 amino acids, each repeat containing two related sub-domains
 CC that flank a pattern of conserved cysteine residues known as an
 CC epidermal growth factor (EGF)-like motif. (I) has neuroprotective,
 CC neotropic and antilipaeamic activities, and can be used as a modulator
 CC of reelin-LDLR interaction. (I) is useful in screen for compounds that
 CC modulate reelin binding to an LDLR, in an assay system, where the assay
 CC system comprises a microplate array and an automated robotic
 CC microprocessor controlled system for adding and removing reagents to
 CC the microplate array. The compounds identified by the above screening
 CC method are useful as therapeutic agents to provide or alleviate a
 CC diverse spectrum of diseases including neurodegenerative disorders such
 CC as Alzheimer's disease, to facilitate neuronal regeneration after

CC injury, to prevent or alleviate lipid metabolism diseases, to enhance
 CC cognitive functions and memory or to ameliorate other developmental
 CC disorders. The present sequence encodes human reelin, which is used in
 CC the exemplification of the present invention.
 XX
 SQ Sequence 11580 BP; 3014 A; 2696 C; 2753 G; 3116 T; 1 other;
 Alignment Scores:
 Pred. No.: 1e-62 Length: 11580
 Score: 572.00 Matches: 106
 Percent Similarity: 95.73% Conservative: 6
 Best Local Similarity: 90.60% Mismatches: 5
 Query Match: 91.52% Indels: 0
 DB: 24 Gaps: 0
 US-09-897-438B-2 (1-117) x ABA92604 (1-11580)
 QY 1 GluGlnCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20
 Db 860 GAACAGTGTGGCGGATTTATGTCATGCAATGCCGTCACTTCTGTGAACCATATGGCCCA 919
 QY 21 ArgGluLeuThrThrThrCysLeuAsnThrThrThrAlaSerValLeuGlnPheSerIle 40
 Db 920 CGAAGAACTGATTAACACAGCCCTTAATACAAACAGCTTCTGCTCCCAATTTTCCATT 979
 QY 41 GlySerGlySerCysArgPheSerTyrSerAspProSerIleThrValSerTyrAlaLys 60
 Db 980 GGGTCAGGTTCATGTCGCTTAGTTATTAGACCCAGCCAGCATCATCTGTATTATGCAAG 1039
 QY 61 AsnAsnThrAlaAspTrpIleGlnLeuGluLysIleArgAlaProSerAsnValSerThr 80
 Db 1040 AATAACTCTGCGACTGGATTACAGTCAGTAGAGAAAATTTAGAGCCCTTCCCAATGTCAGCACCA 1099
 QY 81 ValIleHisIleLeuTyrLeuProGluGluAlaLysGlyGluSerValGlnPheGlnTrp 100
 Db 1100 ATCATCCATATCTCTACCTTCTGAGGACGCCAAAGGGGAGAAATGTCCAATTTTCAGTGG 1159
 QY 101 LysGlnAspSerLeuArgValGlyGluValTyrGluAlaCysTrpAlaLeu 117
 Db 1160 AAGCAGGAAAATCTTCTGTAGGTGAAGTGTATGAAGCCCTGTCGGCCCTTA 1210
 RESULT 7
 ID AAS89484 standard; cdna; 11632 BP.
 XX
 AC AAS89484;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #25288.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 XX
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 XX WPI; 2001-639362/73.
 DR P-PSDB; ABB25297.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity
 XX
 PS Claim 1; SEQ ID No 25288; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 11632 BP; 3028 A; 2716 C; 2770 G; 3117 T; 1 other;

Alignment Scores:

Pred. No.: 1.01e-62 Length: 11632
 Score: 572.00 Matches: 106
 Percent Similarity: 95.73% Conservative: 6
 Best Local Similarity: 90.60% Mismatches: 5
 Query Match: 91.52% Indels: 0
 DB: 23 Gaps: 0

US-09-897-438B-2 (1-117) x AAS99484 (1-11632)

Qy 1 GluInCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20
 Db 860 GAACAGTGTGGCGGATATGCAATGGCAATGCCGTACCTCTCTGTGAACCATATGGCCCA 919
 Qy 21 ArgGluLeuThrThrCysLeuAsnThrThrThrAlaSerValLeuGlnPheSerIle 40
 Db 920 CGAGAACTGATTACACAGGCGCTTAATACAAACACAGCTTCTGCTCCCAATTTCCATT 979
 Qy 41 GlySerGlySerCysArgPheSerTyrSerAspProSerIleThrValSerTyrAlaLys 60
 Db 980 GGGTCAGGTTTCATCGCTTTAGTTATTACAGACCCCGCATCATCGTGTATATGCCAAG 1039
 Qy 61 AsnAsnThrAlaAspTrpIleGlnLeuGluLysIleArgAlaProSerAsnValSerThr 80
 Db 1040 AATAACTCTCGGACTGGATTCAGTCAGTAGAGAAATAGAGCCCTTCCATGTCAGCACA 1099
 Qy 81 ValIleHisIleLeuTyrLeuProGluAlaLysGlyGluSerValGlnPheGlnTrp 100
 Db 1100 ATCATCCATATCTCTCTACCTCTCTGAGGAGCCCAAGGGAGAATGTCCAAATTCAGTGG 1159
 Qy 101 LysGlnAspSerLeuArgValGlyGluValTyrGluAlaCysTrpAlaLeu 117
 Db 1160 AAGCAGAAATCTTCGTGTAGGTGAAGTGTATGAAGCCCTGCTGGGCCCTTA 1210
 RESULT 8
 AAD22779
 ID AAD22779 standard; cDNA; 2025 BP.
 XX
 AC AAD22779;
 XX
 XX 26-FEB-2002 (first entry)
 DT
 XX Xenopus laevis intact reelin cDNA.

XX Frog; reelin; F-spondin domain; CR-50 epitope; gene therapy; agyria;
 KW polymicrogyria; ectopic gray matter; ss.
 OS
 OS Xenopus laevis.

XX Key Location/Qualifiers
 FH 157..2025
 FT /*tag= a
 FT /product= "Xenopus intact reelin protein"
 FT /transl_except= (pos:1606..1608, aa:Cys)
 FT /transl_except= (pos:1639..1641, aa:Cys)
 FT /note= "CDS does not include stop codon"
 FT /partial

PN EPI149844-A2.

PD 31-OCT-2001.

XX 11-APR-2001; 2001EP-0303411.

XX 11-APR-2000; 2000JP-0109954.

XX (RIKE) RIKEN KK.

XX Mikoshiba K, Tabata H, Nakajima K;

XX WPI: 2002-019320/03.

DR P-PSDB; AAE13607.

XX Novel truncated Reelin protein containing F-spondin domain and CR-50
 PT recognition site of Reelin protein, but not having Reelin repeat site,
 PT useful to treat diseases including agyria due to abnormal neuron
 PT alignment

XX Example 1; Fig 1A; 47pp; English.

CC The invention relates to a truncated Reelin protein comprising a
 CC F-spondin domain and a CR-50 recognition site but no reelin protein
 CC repeat site. Reelin is an essential molecule in developing a normal
 CC laminated structure of cerebrium. The truncated reelin protein and its
 CC DNA are useful for treating diseases including agyria, polymicrogyria,
 CC and ectopic gray matter due to abnormal neuronal alignment. Truncated
 CC reelin protein DNA is useful in gene therapy. The present sequence is
 CC a cDNA encoding Xenopus laevis intact reelin protein.

XX Sequence 2025 BP; 541 A; 419 C; 460 G; 604 T; 1 other;

Alignment Scores:

Pred. No.: 2.47e-54 Length: 2025
 Score: 499.00 Matches: 87
 Percent Similarity: 88.89% Conservative: 17
 Best Local Similarity: 74.36% Mismatches: 13
 Query Match: 79.84% Indels: 0
 DB: 24 Gaps: 0

US-09-897-438B-2 (1-117) x AAD22779 (1-2025)

Qy 1 GluInCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20
 Db 847 GAGCAGTGTGGTGCATTTATGCAATGTGTGGGCGAGTCATCTTTGTGATCGGTATGGACCA 906
 Qy 21 ArgGluLeuThrThrCysLeuAsnThrThrThrAlaSerValLeuGlnPheSerIle 40
 Db 907 AGAGAATTGATAACTGTTCAATTAACACACAACTACGGCATCTGTTTTCAGTTCATT 966
 Qy 41 GlySerGlySerCysArgPheSerTyrSerAspProSerIleThrValSerTyrAlaLys 60
 Db 967 GGTCAGGATCGTGCAGGTTTCAGTATTTCAGACCCCTGGAATTTGTGTGTCTATACACAAG 1026
 Qy 61 AsnAsnThrAlaAspTrpIleGlnLeuGluLysIleArgAlaProSerAsnValSerThr 80
 Db 1027 AATAATTCATCTAGTTGGATGCCATTGGAGAGAAATAGTCTCTCTCCAATGTAGCACC 1086

PT brains -
XX Example 4; SEQ ID NO: 7711; 650pp + Sequence Listing; English.
PS
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention.
XX
SQ Sequence 413 BP; 134 A; 86 C; 67 G; 126 T; 0 other;

Alignment Scores:
Pred. No.: 2.74e-09 Length: 413
Score: 144.00 Matches: 36
Percent Similarity: 47.17% Conservative: 14
Best Local Similarity: 33.96% Mismatches: 24
Query Match: 23.04% Indels: 32
DB: 22 Gaps: 4

US-09-897-438B-2 (1-117) x AAK07720 (1-413)
QY 1 GluGlnCysGlyThrIleMethHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20
Db 323 GAACAGTGTGGCGGCGATTATGCATGGCAATGCGCTTCACCTTCTGTGAACCATATGGCCCA 264
QY 21 ArgGluLeuThrThrCysLeuAsnThrThrAlaSerValLeuGlnPheSerile 40
Db 263 CGAGAACTGTGAAGTATGCTATTCATGTACA-----TTATTCATTATAGTATT 213
QY 41 GlySerGlySerCysArgPheSerTyr-----SerAspProSerile 54
Db 212 TACTGAGCTTTTGGCAACAAAAGTTTCTTATGCTCTTAAAGTTCAGATGAAAT--- 156
QY 55 ThrValSerTyrAlaLysAsnThrAlaAspTrpIleGlnLeuGluLysIleArgAla 74
Db 156 ----- 156
QY 75 ProSerAsnValSerThrValIleHisIleLeuTyrLeuProGluGluAlaLysGlyGlu 94
Db 155 ---TCAATATGATTATGAGTCTCTTACAAATATCTT---AAATTTGGGAGGGGTGAT 102
QY 95 SerValGlnPheGlnTrp 100
Db 101 TCTATGGCTATTTTGG 84

RESULT 11
AAK33553/c
ID AAK33553 standard; DNA; 413 BP.
XX
AC AAK33553;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 8110.
XX
DE Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 8110; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention.
XX
SQ Sequence 413 BP; 134 A; 86 C; 67 G; 126 T; 0 other;

Alignment Scores:
Pred. No.: 2.74e-09 Length: 413
Score: 144.00 Matches: 36
Percent Similarity: 47.17% Conservative: 14
Best Local Similarity: 33.96% Mismatches: 24
Query Match: 23.04% Indels: 32
DB: 22 Gaps: 4

US-09-897-438B-2 (1-117) x AAK33553 (1-413)
QY 1 GluGlnCysGlyThrIleMethHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20
Db 323 GAACAGTGTGGCGGCGATTATGCATGGCAATGCGCTTCACCTTCTGTGAACCATATGGCCCA 264
QY 21 ArgGluLeuThrThrCysLeuAsnThrThrAlaSerValLeuGlnPheSerile 40
Db 263 CGAGAACTGTGAAGTATGCTATTCATGTACA-----TTATTCATTATAGTATT 213
QY 41 GlySerGlySerCysArgPheSerTyr-----SerAspProSerile 54
Db 212 TACTGAGCTTTTGGCAACAAAAGTTTCTTATGCTCTTAAAGTTCAGATGAAAT--- 156
QY 55 ThrValSerTyrAlaLysAsnThrAlaAspTrpIleGlnLeuGluLysIleArgAla 74
Db 156 ----- 156
QY 75 ProSerAsnValSerThrValIleHisIleLeuTyrLeuProGluGluAlaLysGlyGlu 94
Db 155 ---TCAATATGATTATGAGTCTCTTACAAATATCTT---AAATTTGGGAGGGGTGAT 102
QY 95 SerValGlnPheGlnTrp 100
Db 101 TCTATGGCTATTTTGG 84

RESULT 12
ABS08422/c
ID ABS08422 standard; DNA; 413 BP.
XX
AC ABS08422;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe from lung SEQ ID No 8413.
XX
DE Human; ds; single exon probe; asthma; lung cancer; COPD; IID;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;

KW pulmonary histiocytosis; lymphangioliomyomatosis; Karagener syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.

OS Homo sapiens.

XX WO200186003-A2.

XX 15-NOV-2001.

XX 30-JAN-2001; 2001WO-US00665.

XX 04-FEB-2000; 2000US-180312P.

XX 26-MAY-2000; 2000US-207456P.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-234687P.

XX 27-SEP-2000; 2000US-236359P.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2002-114183/15.

XX Spatially-addressable set of single exon nucleic acid probes, used to

XX measure gene expression in human lung samples -

XX Claim 1; SEQ ID No 8413; 634pp; English.

XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a single exon
 CC probe of the invention.

XX Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 413 BP; 134 A; 86 C; 67 G; 126 T; 0 other;

Alignment Scores: 2.74e-09 Length: 413
 Pred. No.: 144.00 Matches: 36
 Percent Similarity: 47.17% Conservative: 14
 Best Local Similarity: 33.96% Mismatches: 24
 Query Match: 23.04% Indels: 32
 DB: 24 Gaps: 4

US-09-897-438B-2 (1-117) x ABS08422 (1-413)

Qy 1 GluGlnCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20

Db 323 GAACAGTGTGGCGCATATATGCGCAATGCCGTCACCTCTGTGAACCATATGGCCCA 264

Qy 21 ArgGluLeuThrThrCysLeuAsnThrThrAlaSerValLeuGlnPheSerIle 40

Db 263 CGAGAACTGGTAAGTATGCTATTCATGTACA-----TTATTCCATTATAGTATT 213

Qy 41 GlySerGlySerCysArgPheSerTyr-----SerAspProSerIle 54

Db 212 TACTGAGCTTTTGGCAACAAAGTTTTCCTTATGCTCTTAAAGTTCAGATGAAAAT--- 156

Qy 55 ThrValSerTyrAlaLysAsnThrAlaAspTrpIleGlnLeuGluLysIleArgAla 74

Db 156 ----- 156

Qy 75 ProSerAsnValSerThrValIleHisIleLeuTyrLeuProGluAlaLysGlyGlu 94

Db 155 ---TCAAATATGATTATGAGTCCTATTACAAATATCTT---AAAAATGGGAGGGGTGAT 102

Qy 95 SerValGlnPheGlnTrp 100

Db 101 TCTATGGCTATTTTGG 84

RESULT 13

AAS89475

ID AAS89475 standard; cDNA; 4576 BP.

AC AAS89475;

XX 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #25279.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG25288.

XX New isolated polynucleotide and encoded polypeptides, useful in

XX diagnostics, forensics, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits and to assess

XX biodiversity -

XX Claim 1; SEQ ID No 25279; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 4576 BP; 1222 A; 1023 C; 1109 G; 1222 T; 0 other;

Alignment Scores:
Pred. No.: 0.00246 Length: 4576
Score: 109.50 Matches: 37
Percent Similarity: 48.36% Conservative: 22
Best Local Similarity: 30.33% Mismatches: 52
Query Match: 17.52% Indels: 11
DB: 23 Gaps: 5

US-09-897-438B-2 (1-117) x AAS89475 (1-4576)

Qy 3 CysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGlyProArgGlu 22
Db 2855 TGTGGAACACTTAACGATGGCAATCTCTACTTCAATGGCCCT---GGGAAAGGGAA 2911
Qy 23 LeuThrThrCysLeuAsnThrThrAlaSerValLeuGlnPhe-----SerIleThr 40
Db 2912 GCCCGAGCGTCCCTCTGGACACAGGAAATATCAGACTTGTCAATTTATATACAAATT 2971
Qy 41 GlySerGlySerCysArgPheSerTyrSerAspPro-----SerIleThr 55
Db 2972 GGAACAAACTTCAGGCATTACCTGCATCAACCAAGAACTAGAAATGAGGCTTATT 3031
Qy 56 ValSerTyrAlaLysAsnThrAlaAspTrpIleGlnLeuGluLysIleArgAlaPro 75
Db 3032 GTTCAGTATTCAATGACAATGGGATCTCTGGCATTTGCTTCGAGAGTTGGACTTCATG 3091
Qy 76 SerAsnValSerThrValIleHisIleLeuTyrLeuProGluGluAlaLysGlyCysSer 95
Db 3092 TCCCTTCGGAAACACACATCATTTCCATTGACCTGGCCAGGACGGGAAAGACACCTGCA 3151
Qy 96 ValGlnPheGlnTrpLysGlnAspSerLeuArgValGlyGluValTyrGluAlaCysTrp 115
Db 3152 ACGGCATTTCGATGGTGCACCG-----CAACATGGGAG---CATTGACCCCATGG 3202
Qy 116 AlaLeu 117
Db 3203 GCCTTG 3208
RESULT 14
AAK20384/C
ID AAK20384 standard; DNA; 87 BP.
XX AC AAK20384;
XX AC AAK20384;
XX 05-NOV-2001 (first entry)
DE Human brain expressed single exon probe SEQ ID NO: 20375.
XX

KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer; ss.
XX Homo sapiens.
XX OS
XX WO200157275-A2.
XX PD 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US00667.
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0623366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX Example 4; SEQ ID NO: 20375; 650pp + Sequence Listing; English.
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is one of the probes of the
XX invention.
XX Sequence 87 BP; 21 A; 25 C; 19 G; 22 T; 0 other;

Alignment Scores:
Pred. No.: 1.88e-05 Length: 87
Score: 107.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 17.12% Indels: 0
DB: 22 Gaps: 0

US-09-897-438B-2 (1-117) x AAK20384 (1-87)

Qy 1 GluGlnCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGly 19
Db 59 GAACAGTGTGGCGGATATGTCATGGCAATGCCCTCCTCTGTGAACCATATGGC 3
RESULT 15
AAK46490/C
ID AAK46490 standard; DNA; 87 BP.
XX AC AAK46490;
XX 06-NOV-2001 (first entry)
XX Human bone marrow expressed single exon probe SEQ ID NO: 21047.
XX Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX Homo sapiens.
XX OS
XX WO200157276-A2.
XX PD 09-AUG-2001.

```

XX
PF 30-JAN-2001; 2001WO-US00658.
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 21047; 658pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention.
XX
XX SQ Sequence 87 BP; 21 A; 25 C; 19 G; 22 T; 0 other;

Alignment Scores:
Pred. No.: 1.88e-05 Length: 87
Score: 107.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 17.12% Indels: 0
DB: 22 Gaps: 0

US-09-897-438B-2 (1-117) x AAK46490 (1-87)

QY 1 GluGlnCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGly 19
Db 59 GAACAGTGTGGCGGATTATGTCATGGCAATGCCGTCACCTTCTGTGAACCATATGGC 3

Search completed: November 6, 2002, 19:30:02
Job time : 308 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 6, 2002, 20:18:37 ; Search time 87 Seconds
(without alignments)
448.036 Million cell updates/sec

Title: US-09-897-438B-2

Perfect score: 625

Sequence: 1 ECGTTHGNAVTFCEPYGP.....FQWKQSLRVGEVYACWAL 117

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 310279 seqs, 166577418 residues

Total number of hits satisfying chosen parameters: 620558

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlp
-Q=/cgn2_1/US09897438/runat_06112002_101927_11396/app_query.fasta_1.263
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09897438 -CGN_1_1_36 @runat_06112002_101927_11396
-NCPU=6 -ICPU=3 -NO_XLPY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	625	100.0	351	10	US-09-897-438B-1
2	572	91.5	11580	10	US-09-880-107-3436
3	224	35.8	499	10	US-09-864-761-14497
4	190	30.4	214	10	US-09-864-761-31037

c 5	144	23.0	413	10	US-09-864-761-11808	Sequence 11808, A
c 6	107	17.1	87	10	US-09-864-761-28379	Sequence 28379, A
c 7	72	11.5	795	10	US-09-974-300-2745	Sequence 2745, Ap
c 8	70	11.2	9274	10	US-09-885-535-3	Sequence 3, Appli
c 9	69	11.0	495	10	US-09-864-761-10582	Sequence 10582, A
c 10	68.5	11.0	468	10	US-09-917-800A-547	Sequence 547, App
c 11	68	10.9	978	10	US-09-886-055-2	Sequence 2, Appli
c 12	67	10.7	1017	10	US-09-802-371-3	Sequence 3, Appli
c 13	67	10.7	1585	10	US-09-802-371-1	Sequence 1, Appli
c 14	66.5	10.6	612	10	US-09-974-300-1207	Sequence 1207, Ap
c 15	66.5	10.6	2082	10	US-09-741-669-200	Sequence 200, App
c 16	66.5	10.6	143058	10	US-09-967-769A-316	Sequence 316, App
c 17	66	10.6	2001	10	US-09-841-683-3	Sequence 3, Appli
c 18	66	10.6	2466	10	US-09-922-138-4	Sequence 4, Appli
c 19	66	10.6	4237	10	US-09-764-877-3835	Sequence 3835, Ap
c 20	66	10.6	5504	8	US-08-913-322-1	Sequence 1, Appli
c 21	66	10.6	6124	8	US-08-913-322-21	Sequence 21, Appli
c 22	66	10.6	6124	10	US-09-967-769A-184	Sequence 184, App
c 23	66	10.6	6133	8	US-08-913-322-2	Sequence 2, Appli
c 24	66	10.6	6228	8	US-08-913-322-23	Sequence 23, Appli
c 25	65.5	10.5	2846	10	US-09-962-436-303	Sequence 303, App
c 26	65.3	10.5	2846	10	US-09-962-436-309	Sequence 309, App
c 27	65.5	10.5	5769	10	US-09-964-824A-284	Sequence 284, App
c 28	65	10.4	528	10	US-09-974-300-6115	Sequence 6115, Ap
c 29	65	10.4	582	10	US-09-974-300-6100	Sequence 6100, Ap
c 30	65	10.4	954	12	US-10-043-238-2	Sequence 2, Appli
c 31	64.5	10.3	722	10	US-09-925-300-608	Sequence 608, App
c 32	64.5	10.3	1653	10	US-09-815-242-6402	Sequence 6402, Ap
c 33	64.5	10.3	2073	10	US-09-815-242-8362	Sequence 8362, Ap
c 34	64.5	10.3	2109	10	US-09-815-242-4555	Sequence 4555, Ap
c 35	64	10.2	373	10	US-09-783-590-5268	Sequence 5268, Ap
c 36	64	10.2	1341	10	US-09-350-756-2	Sequence 2, Appli
c 37	64	10.2	1740	10	US-09-974-300-479	Sequence 479, App
c 38	64	10.2	4104	10	US-09-801-368-107	Sequence 107, App
c 39	64	10.2	4487	10	US-09-738-968-32	Sequence 32, Appli
c 40	63.5	10.2	670	10	US-09-735-705-7	Sequence 7, Appli
c 41	63.5	10.2	670	10	US-09-850-716A-7	Sequence 7, Appli
c 42	63.5	10.2	670	10	US-09-897-778-7	Sequence 7, Appli
c 43	63.5	10.2	908	12	US-10-044-090-213	Sequence 213, App
c 44	63.5	10.2	1678	10	US-09-925-300-153	Sequence 153, App
c 45	63.5	10.2	1830	10	US-09-815-242-7661	Sequence 7661, Ap

ALIGNMENTS

RESULT 1
US-09-897-438B-1
; Sequence 1, Application US/09897438B
; Patent No. US20020137095A1
; GENERAL INFORMATION:
; APPLICANT: Mikoshiba, Katsuhiko
; APPLICANT: Tate, Naoko
; TITLE OF INVENTION: REELIN PROTEIN CR-50 EPTOPE REGION
; FILE REFERENCE: 04853-0076-00000
; CURRENT APPLICATION NUMBER: US/09/897,438B
; CURRENT FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: JP 2000-202801
; PRIOR FILING DATE: 2000-07-04
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 351
; TYPE: DNA
; ORGANISM: Mus musculus
; US-09-897-438B-1

Alignment Scores:
Pred. No.: 4.49e-79
Score: 625.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 10
Length: 351
Matches: 117
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-897-438B-2 (1-117) x US-09-897-438B-1 (1-351)

Qy 1 GluGlnCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20
 Db 1 GAGCAGTGTGGCACCATCATGCGATGGCAATGCTGTACCTTCTGTGAGCGGTACGGCCCT 60
 Qy 21 ArgGluLeuThrThrThrCysLeuAsnThrThrThrAlaSerValLeuGlnPheSerIle 40
 Db 61 CGAGAGCTGACCACCATCCCTGAACACAAACAGCATCTGCTCTCCAGTTTCCATT 120
 Qy 41 GlySerGlySerCysArgPheSerTyrSerAspProSerIleThrValSerTyrAlaLys 60
 Db 121 GGGTCAGGATCATGTGATTTAGTTACTCTGACCCAGCATCACTGTGTCATACGCCAAG 180
 Qy 61 AsnAsnThrAlaAspTrpIleGlnLeuGluLysIleArgAlaProSerAsnValSerThr 80
 Db 181 AACAAATACCGCTGATGGATTACGTGGAGAAATTTAGAGCCCTTCCAAATGTGAGCACA 240
 Qy 81 ValIleHisIleLeuTyrLeuProGluGluAlaLysGlyGluSerValGlnPheGlnTrp 100
 Db 241 GTCATCCACATCCTGTACCTCCCGGAGGAAGCCAAAGGGAGAGCGTGCAGTTCCAGTGG 300
 Qy 101 LysGlnAspSerLeuArgValGlyGluValTyrGluAlaCysTrpAlaLeu 117
 Db 301 AAACAGACAGCCTGCGAGTGGTGAGGTGATGAGGCCTGTGGGCCCTG 351

RESULT 2

US-09-880-107-3436

; Sequence 3436, Application US/09880107

; Patent No. US20020142981A1

; GENERAL INFORMATION:

; APPLICANT: Horne, Darci T.

; APPLICANT: Vockley, Joseph G.

; APPLICANT: Scherf, Uwe

; APPLICANT: Gene Logic, Inc.

; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer

; FILE REFERENCE: 44921-5028-WO

; CURRENT APPLICATION NUMBER: US/09/880,107

; CURRENT FILING DATE: 2001-06-14

; PRIOR APPLICATION NUMBER: US 60/211,379

; PRIOR FILING DATE: 2000-06-14

; PRIOR APPLICATION NUMBER: US 60/237,054

; PRIOR FILING DATE: 2000-10-02

; NUMBER OF SEQ ID NOS: 3950

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3436

; LENGTH: 11580

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U97916

US-09-880-107-3436

Alignment Scores:

Pred. No.:	2,32e-69	Length:	11580
Score:	572.00	Matches:	106
Percent Similarity:	95.73%	Conservative:	5
Best Local Similarity:	90.60%	Mismatches:	6
Query Match:	91.52%	Indels:	0
DB:	10	Gaps:	0

US-09-897-438B-2 (1-117) x US-09-880-107-3436 (1-11580)

Qy 1 GluGlnCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20
 Db 860 GAACAGTGTGGCCGATATGATGCAATGCGTCACCTTCTGTGAACCATATGGCCCA 919
 Qy 21 ArgGluLeuThrThrThrCysLeuAsnThrThrThrAlaSerValLeuGlnPheSerIle 40
 Db 920 CGAAGATGATTTACCAACGCGCTTAATACAAACAGCTTCTGTCTCCAAATTTCCATT 979
 Qy 41 GlySerGlySerCysArgPheSerTyrSerAspProSerIleThrValSerTyrAlaLys 60

```

Db      980  GGGTCAGGTTTCATGTCGCTTTTATTAGTATTAGAACCCAGCATCATCGTGTATATGCCAAG 10399
Qy      61  AsnAsnThrAlaAspTrpIleGlnLeuGluLysIleArgAlaProSerAsnValSerThr 80
Db     1040  AATAACTCTCGGCACTGGATTTCAGCTAGAGAAAATTAGAGCCCTTCCAATGTCAGCACA 10999
Qy      81  ValIleHisIleLeuTyrLeuProGluGluAlaLysGlyGluSerValGlnPheGlnTrp 100
Db     1100  ATCATCATATCTCTACCTTCCTGAGGAGCGCAAGGGAGAAATGTCCAATTTTCAGTGG 11599
Qy     101  LysGlnAspSerLeuArgValGlyGluValTyrGluAlaCysTrpAlaLeu 117
Db     1160  AAGCAGGAAAATCTTCGTAGGTGAAGTGTATGAAGCCTGCTGGGCCCTTA 1210

RESULT 3
US-09-864-761-14497/c
; Sequence 14497, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

```

```

RESULT 3
US-09-864-761-14497/c
; Sequence 14497, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED
; TITLE OF INVENTION: GENE EXPRESSION ANAL
; FILE REFERENCE: Aomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine
; SEQ ID NO 14497
; LENGTH: 499
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC000121.1

```

```

% NUMBER OF SEQ ID NOS: 4911
% SOFTWARE: Annonax Sequence Listing Engine ver. 1.1.1
% SEQ ID NO 14497
% LENGTH: 499
% TYPE: DNA
% ORGANISM: Homo sapiens
% FEATURE:
% OTHER INFORMATION: MAP TO AC000121.1

```



```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 11808
; LENGTH: 413
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC000121.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.55
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.57
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.51
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.44
US-09-864-761-11808

Alignment Scores:
Pred. No.: 1.95e-11 Length: 413
Score: 144.00 Matches: 36
Percent Similarity: 47.17% Conservative: 14
Best Local Similarity: 33.96% Mismatches: 24
Query Match: 23.04% Indels: 32
DB: 10 Gaps: 4

US-09-897-438B-2 (1-117) x US-09-864-761-11808 (1-413)

Qy 1 GluGlnCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
Db 323 GAACAGTGTGGCGGATGATTGATGCGAATGCGGTCACCTTCTGTGAACCATATGCGCCA 264

Qy 21 ArgGluLeuThrThrCysLeuAsnThrThrAlaSerValLeuGlnPheSerIle 40
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
Db 263 CGAACTGGTAGTGTGCTATTCATGATACA-----TTATTCATATAGTATT 213

Qy 41 GlySerGlySerCysArgPheSerTyr-----SerAspProSerIle 54
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
Db 212 TACTGAGCTTTTCCCAACAAAGTTTCTTATGCTCTTAAAGTTCAGATGAAAT--- 156

Qy 55 ThrValSerTyrAlaLysAsnAsnThrAlaAspTrpIleGlnLeuGluLysIleArgAla 74
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
Db 156 ----- 156

Qy 75 ProSerAsnValSerThrValIleHisIleLeuTyrLeuProGluGluAlaLysGlyGlu 94
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
Db 155 ---TCAATATGATTATGAGTCTTATCAAAATATCTT---AAATTTGGAGGGGTAT 102

Qy 95 SerValGlnPheGlnTrp 100
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
Db 101 TCTATGCGCTATTTTGG 84

RESULT 6
US-09-864-761-28379/c
; Sequence 28379, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
```

```
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28379
; LENGTH: 87
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC000121.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.55
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.57
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.51
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.44
; OTHER INFORMATION: SWISSPROT HIT: P76008, EVALUE 3.20e+00
; OTHER INFORMATION: NT HIT: g14826977, EVALUE 5.00e-42
US-09-864-761-28379

Alignment Scores:
Pred. No.: 2.98e-07 Length: 87
Score: 107.00 Matches: 18
Percent Similarity: 94.74% Conservative: 0
Best Local Similarity: 94.74% Mismatches: 1
Query Match: 17.12% Indels: 0
DB: 10 Gaps: 0

US-09-897-438B-2 (1-117) x US-09-864-761-28379 (1-87)

Qy 1 GluGlnCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGly 19
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
Db 59 GAACAGTGTGGCGGATGATTGATGCGAATGCGGTCACCTTCTGTGAACCATATGCGC 3

RESULT 7
US-09-974-300-2745/c
; Sequence 2745, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
```

; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE OF INVENTION: Expression
; FILE REFERENCE: 10085,500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2745
; LENGTH: 795
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-2745

Alignment Scores:
Pred. No.: 0.669 Length: 795
Score: 72.00 Matches: 22
Percent Similarity: 41.89% Conservative: 9
Best Local Similarity: 29.73% Mismatches: 43
Query Match: 11.52% Indels: 0
DB: 10 Gaps: 0

US-09-897-438B-2 (1-117) x US-09-974-300-2745 (1-795)

Qy 32 ThrAlaSerValLeuGlnPheSerIleGlySerCysArgPheSerTyrSerAsp 51
|||||
Db 360 ACGGCTTCTAGCGCAATGTCAACACGAGCGGCGCTTTCGTGAGATTAAAG 301
Qy 52 ProSerIleThrValSerTyrAlaLysAsnThrAlaAspTrpIleGlnLeuGluLys 71
|||
Db 300 CCGTATTGGCCGATCCTGCCGCCATGCCAGACAGGTGATGATAAATGGCCGCAAAA 241
Qy 72 IleArgAlaProSerAsnValSerThrValIleHisIleLeuTyrLeuProGluAla 91
|||
Db 240 ACAGCGAGCCGATGCGCAGCAGCAAAATCAGCCATATTTGATTCGCCAGCCCTTGAAGCA 181
Qy 92 LysGlyGluSerValGlnPheGlnTrpLysGlnAspSerLeu 105
|||||
Db 180 ACAGTCGACTCGATAATCAAGTTGTCGCAAAACGACAGCATA 139

RESULT 8
US-09-885-535-3/C
; Sequence 3, Application US/09885535
; Patent No. US20020104105A1
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Heichman, Karen
; APPLICANT: Bartel, Paul L.
; TITLE OF INVENTION: Protein-Protein Interactions
; FILE REFERENCE: 2318-266-II
; CURRENT APPLICATION NUMBER: US/09/885,535
; CURRENT FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/213,245
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 9274
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-885-535-3

Alignment Scores:
Pred. No.: 50.5 Length: 9274
Score: 70.00 Matches: 24
Percent Similarity: 59.09% Conservative: 15
Best Local Similarity: 36.36% Mismatches: 17
Query Match: 11.20% Indels: 10
DB: 10 Gaps: 4

US-09-897-438B-2 (1-117) x US-09-885-535-3 (1-9274)
Qy 24 ThrThrThrCysLeuAsnThrThrAlaSerValLeuGlnPheSerIleGlySer--- 42
|||||
Db 5024 AGCACATGATGCTTG-----GCAAGTTCAATTTCAATGGCTTTAGGCTCTCCT 4977
Qy 43 -----GlySerCysArgPheSerTyrSerAspProSerIleThrValSerTyrAlaLys 60
|||
Db 4976 CCAACAGGTTTCTGC-----TCACTTAGCAAGCCCTCGGTGTGTGTCAGCCATGCCAGG 4923
Qy 61 AsnAsnThr---AlaAspTrpIleGlnLeuGluLysIleArgAlaProSerAsnValSer 79
|||||
Db 4922 AGCTCATCCAGGCATGTTGGAAGTACCAAGGCTAATAGACACCCCTCCAGTTTATGC 4863
Qy 80 ThrValIleHisIleLeu 85
|||||
Db 4862 TGTCTGTTGATGATCTC 4845

RESULT 9
US-09-864-761-10582/C
; Sequence 10582, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Rank, Sharon G.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL F
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 10582
; LENGTH: 495
; TYPE: DNA


```

QY 2 GlnCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGlyProArg 21
Db 338 AGATGTACTATTTCTATTTGTGTGTGTCATTCAGCAATTTGCTCT---TGGGGACCATGG 394
QY 22 GluLeuThrThrCysLeuAsnThrThrAlaSerValLeuGlnPheSerIleGly 41
Db 395 CTTATG---ACCATTGTTGGGAGCTGCCACCCCTCTGAATTTATACAATTCATGCGGC 451
QY 42 SerGlySer----- 44
Db 452 CCAGGTTGGGCATTTTGCTCACAGTCATCTCATGGTTCCTCAGTAATATTATTGCTCTGA 511
QY 45 -----CysArgPheSerTyrSerAspProSerIleThrValSerTyrAlaLys 60
Db 512 CACACACCCCTTCGTCATTCATTAATTCCTTCTGTGAACCACCAACACTCTCCACACACTTCT 571
QY 61 AsnAsnThrAlaAsp----- 65
Db 572 TCCTGACTGGCCCTCTGCTCAAACTGCTCTGTTTCAGATCATGTGATCAATGAGCTTG 631
QY 66 -----TrpIle-----GlnLeuGluLysIleArgAlaPro 75
Db 632 TGTGTTTATTGGGGTTTATCAGTTATCATCTTCCCTTTACACTCAGCTTCCTTTTCTCT 691
QY 76 SerAsnValSerThrValIleHisIleLeuTyrLeuProGluGluAlaLysGlyGlu 94
Db 692 ATGCTGTCATCATCAGAGCTGCTCAGAGTATCTTCCACACAGGAAAGTGGAAAG 748

RESULT 12
US-09-802-371-3
; Sequence 3, Application US/09802371
; Patent No. US20010036649A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Rudolph-Owen, Laura
; TITLE OF INVENTION: 26934, A No. US20010036649A1el Cytidine Deaminase-Like
; FILE REFERENCE: 35800/213921
; CURRENT APPLICATION NUMBER: US/09/802,371
; CURRENT FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/188,294
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1017
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-802-371-3

Alignment Scores:
Pred. No.: 4.87 Length: 1017
Score: 67.00 Matches: 25
Percent Similarity: 44.66% Conservative: 21
Best Local Similarity: 24.27% Mismatches: 39
Query Match: 10.72% Indels: 18
DB: 10 Gaps: 4

US-09-897-438B-2 (1-117) x US-09-802-371-3 (1-1017)

QY 2 GlnCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGlyProArg 21
Db 334 CAGATTGCTCTTATTAAACATGGTCAAGGCTCAAAACATGTGATCTTTATTTTCCAGA 393
QY 22 GluLeuThrThrCysLeuAsnThrThrAlaSerValLeuGlnPheSerIleGly 41
Db 394 AAACCATGTTCTCTCTGT-----TTGAAATGATTGTAAT 429
QY 42 SerGlySerCysArgPheSerTyr-----SerAspProSerIleThrVal----- 56
Db 430 GCTGGAGTTAACCGAATTCCTACTGCGCTGCTGATGAGAAATGAAGTTTGGTTACGGAG 489
QY 57 ---SerTyrAlaLysAsnAsnThrAlaAspTrpIleGlnLeuGluLysIleArgAlaPro 75

```

RESULT 14

US-09-974-300-1207/c
; Sequence 1207, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods for Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1207
; LENGTH: 612
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-1207

Alignment Scores:

Pred. No.:	2.68	Length:	612
Score:	66.50	Matches:	26
Percent Similarity:	40.59%	Conservative:	15
Best Local Similarity:	25.74%	Mismatches:	44
Query Match:	10.64%	Indels:	16
DB:	10	Gaps:	2

US-09-897-438B-2 (1-117) x US-09-974-300-1207 (1-612)

Qy	3	CysGlyThrIleMethHisGlyAsnAlaValThrPheCysGluProTyrglyProArgGlu	22
Db	289	TGCTCGCGCTTAAATGTTTCGGCGCGTATGTCGCTTTCTGT	251
Qy	23	LeuThrThrCysLeuAsnThrThr-----AlaSerValLeu	36
Db	250	-----TCATCCTGTTTTCGGGGTTACCGAACGTCATGATTCGGAAACGTTTCG	197
Qy	37	GlnPheSerIleGlySerGlySerCysArgPheSerTyrrSerAspProSerIleThrVal	56
Db	196	GAAGAAGCGCAGCTCCGGCCCTTGTCAGCTTATCGTGGAGCGTGCGACATTATCTGTT	137
Qy	57	SerTyrrAlaLysAsnAsnThrAlaAspTrpIle-GlnLeuGluLysIleArgAlaProSe	76
Db	136	TCTTCGATGTCACCTCAACCTTGGCGAGCATATCGCCTGCATCCAAFTTTTCGACCATG	77
Qy	76	rAsnValSerThrValIleHisIleLeuTyrrLeuProGluGluAlaLysGlyGluSerVa	96
Db	76	TACATGATCGTCACGCTGTTTCTCTTTCCTCCAGATGGCGTAAATGATCGGGGAT	17
Qy	96	1 96	
Db	16	C 16	

RESULT 15

US-09-741-669-200/c
; Sequence 200, Application US/09741669
; Patent No. US2002022718A1
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; TITLE OF INVENTION: Genes identified as required for
; TITLE OF INVENTION: proliferation of E. coli
; FILE REFERENCE: ELITRA.009A
; CURRENT APPLICATION NUMBER: US/09/741,669
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/173005
; PRIOR FILING DATE: 1999-12-23

; NUMBER OF SEQ ID NOS: 481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 200
; LENGTH: 2082
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2082)
US-09-741-669-200

Alignment Scores:	16.7	Length:	2082
Pred. No.:	66.50	Matches:	25
Score:	44.32%	Conservative:	14
Percent Similarity:	28.41%	Mismatches:	42
Best Local Similarity:	10.64%	Indels:	7
Query Match:	10	Gaps:	4

US-09-897-438B-2 (1-117) x US-09-741-669-200 (1-2082)

Qy	2	GlnCysGlyThrIleMethHisGlyAsnAlaValThrPheCysGluProTyrgly---	20
Db	1637	AGGTGTAGCTCACCTTGTTTAAACGACGCCATCACCGCTGTTCTCGGCGAGTTTCATC	1578
Qy	21	ArgGluLeuThrThrCysLeuAsnThrThrAlaSerValLeuGlnPheSerIle	40
Db	1577	CGCCCGTGTACACGCGCAACGTTCAACTCTGTAGCGGCAGTTTCAACTCTTCCAGGTA	1518
Qy	41	GlySerGlySerCysArgPheSerTyrrSerAspProSerIleThrVal-----	58
Db	1517	GCTTCCGCGCGCTGCGCTTCCAGCAATTCGACTCTTCAATCAACGTACAAACCCAGTAT	1458
Qy	59	Ala-----LysAsnAsnThrAlaAspTrpIle-----GlnLeuGluLysIleArgAla	74
Db	1457	GCCTGACGACCTTCAGTTATGCGAGCGGTGTCACGCGGTCAATGATGTCGGTACGCGG	1398
Qy	75	ProSerAsnValSerThrValIle	82
Db	1397	GTATCAGGAATAGCGACCGTAGTC	1374

Search completed: November 6, 2002, 22:06:36
Job time : 91 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 6, 2002, 19:22:22 ; Search time 2216 Seconds
(without alignments)
855.087 Million cell updates/sec

Title: US-09-897-438B-2

Perfect score: 625

Sequence: 1 EQCGTMRHGNVATFCEPYGP.....FQWKQDSLRLVGEVYEAACWAL 117

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL-frame_p2n.model -DEV-xlp
-Q/cgn2_1/USPTO.spool/US09897438/runat_06112002_101924_11301/app_query.fasta_1.263
-DB-EST -QFMT-fastap -SUFFIX=est -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX-blosum62 -TRANS-human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT-ptc -NORM-ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER-US09897438 @CGN_1_1.1716 @runat_06112002_101924_11301 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estnu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_gss.*
18: em_gss_hum.*
19: em_gss_inv.*
20: em_gss_pln.*
21: em_gss_vrt.*
22: em_gss_fun.*
23: em_gss_mam.*
24: em_gss_mus.*
25: em_gss_other.*
26: em_gss_pro.*
27: em_gss_rod.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
C 1	437	69.9	315	12	BE938667 QV0-TN008
C 2	401	64.2	600	12	BG803882 0243-03 M
C 3	164	26.2	635	10	BB248113 BB248113
C 4	132	21.1	568	12	BF387628 UI-R-CAL-
C 5	120.5	19.3	1022	17	AL174752 Tetraodon
C 6	96	15.4	979	14	BQ942386 AGENCOURT
C 7	92	14.7	743	10	BB502922 BB502922
C 8	87.5	14.0	709	14	BQ444265 UI-M-EXO-
C 9	86.5	13.8	670	10	BB628782 BB628782
C 10	86.5	13.8	745	14	BQ180824 UI-M-EXO-
C 11	85	13.6	942	12	BB92165 601434685
C 12	84	13.4	369	10	AW158822 za45g06.X
C 13	84	13.4	674	13	BI739930 603362473
C 14	84	13.4	692	14	BQ769321 UI-M-FIO-
C 15	84	13.4	706	14	BQ771419 UI-M-FIO-
C 16	83	13.3	594	10	BE361240 DGL-70.B0
C 17	82.5	13.2	719	14	BQ119839 EST605415
C 18	82.5	13.2	1458	11	AK017094 Mus muscu
C 19	81.5	13.0	406	13	BI701324 sag57d09.
C 20	81.5	13.0	418	12	BF598079 sv03b03.Y
C 21	81.5	13.0	420	13	BI700401 sag60d05.
C 22	81.5	13.0	421	14	BQ473635 sap15g06.
C 23	81.5	13.0	449	13	BI700180 sag64d10.
C 24	81.5	13.0	529	13	BM307196 sak38a07.
C 25	81.5	12.9	546	10	BE440529 sp46f03.Y
C 26	80.5	12.9	655	13	BM598586 170006876
C 27	80	12.8	455	10	AW036062 EST278076
C 28	80	12.8	631	14	BM861182 fy46h06.Y
C 29	80	12.8	742	17	AG115800 Pan trogl
C 30	79.5	12.7	662	12	BF211911 601813083
C 31	79	12.6	1701	13	BM467416 AGENCOURT
C 32	78.5	12.6	603	9	AI386066 mm17c06.Y
C 33	78.5	12.6	665	10	AV930527 AV930527
C 34	78.5	12.6	936	14	BQ82577 AGENCOURT
C 35	78	12.5	511	12	BG378762 UI-R-CV1-
C 36	77.5	12.4	477	14	M75904 CEL06C4 Chr
C 37	77.5	12.4	639	17	AZ056749 RPCI-23-4
C 38	77.5	12.4	645	13	BJ501604 BJ501604
C 39	77	12.3	456	9	AA943462 EST198961
C 40	77	12.3	539	10	AW141286 EST291327
C 41	77	12.3	594	10	AW061669 660012H07
C 42	77	12.3	671	17	AG074894 Pan trogl
C 43	77	12.3	831	10	BE535985 601062413
C 44	76.5	12.2	501	13	BJ220506 BJ220506
C 45	76.5	12.2	723	13	BJ320807 BJ320807

ALIGNMENTS

RESULT 1 BE938667/c 315 bp mRNA linear EST 02-OCT-2000
LOCUS QV0-TN0084-180800-342-a08 TN0084 Homo sapiens cDNA, mRNA sequence.
DEFINITION BE938667
ACCESSION BE938667
VERSION EST.
KEYWORDS GI:10466224
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 315)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=5&t2=QV0-TN0084-180>)
800-342-a08&t3=2000-08-18&t4=1
Seq primer: puc 18 forward
High quality sequence stop: 315.

```

FEATURES
  source
    high quality sequence stop: 315.
    Location/Qualifiers
      1..315
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone_lib="TN0084"
        /dev_stage="Adult"
        /note="Organ: Testis; normal; Vector: puc18; Site_1: SmaI;
        Site_2: SmaI; A mini-library was made by cloning products
        derived from ORESTES PCR (U.S. Letters Patent application
        No. 196,716 - Ludwig Institute for Cancer Research)
        profiles into the pUC 18 vector. Reverse transcription of
        tissue mRNA and cDNA amplification were performed under
        low stringency conditions."
      84..64 c 78 g 89 t

```

Alignment Scores:			
Pred. No.:	1..43E-45	Length:	315
Score:	437.00	Matches:	82
Percent Similarity:	96.67%	Conservative:	5
Best Local Similarity:	91.11%	Mismatches:	3
Query Match:	69.92%	Indels:	0
DB:	12	Gaps:	0
US-09-897-438B-2 (1-117) x BE938667 (1-315)			
QY	28	LeuAsnThrThrAlaSerValLeuGlnPheSerIleGlySerGlySerCysArgPhe	47
DB	304	CITAAATACAAACACGCTCTGCTCTCCAAATTTTCATTTGGGTCAAGTTCATGTCGCTTT	245
QY	48	SerTyrSerAspProSerIleThrValSerTyrAlaLysAsnAsnThrAlaAspTrpIle	67
DB	244	AGTTATTACAGCCCCAGCATCATCGTGTTATATGCCAGAATAACTCTGCGGACATGGATT	185
QY	68	GlnLeuGluLysIleArgAlaProSerAsnValSerThrValIleHisIleLeuTyrLeu	87
DB	184	CAGCTAGAGAAAATATAGAGCCCTTCCCAATGTCAGACAAATCATCATATCCTCTACCTT	125
QY	88	ProGluGluAlaLysGlyGluSerValGlnPheGlnTrpLysGlnAspSerLeuArgVal	107
DB	124	CCTGAGGACGCCAAGGGGAGAATTCCAAATTCAGTGGAGCAGGAAGATCTTCGTGTA	65
QY	108	GlyGluValTyrGluAlaCysTrpAlaLeu	117
DB	64	GCTGAAGTGATGAAGCCCTGCTGGCCCTTA	35

RESULT 2					
BG803882/c					
LOCUS	600 bp	mRNA	linear	EST 20-DEC-2001	
DEFINITION	0243-03 Mouse E14.5 retina	lambda ZAP II Library	Mus musculus	CDNA,	
				mRNA sequence.	
ACCESSION	BG803882				

VERSION	EG803882.1	GI:17950794
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE	1 (bases 1 to 600)	
AUTHORS	Mu, X., Zhao, S., Pershad, R., Hsieh, T.-F., Scarpa, A., Wang, S.W., White, R.A., Beremand, P.D., Thomas, T.L., Gan, L. and Klein, W.H.	
TITLE	Gene expression in the developing mouse retina by EST sequencing and microarray analysis	
JOURNAL	Nucleic Acids Res. 29 (24), 4983-4993 (2001)	
MEDLINE	21671825	
COMMENT	Contact: Klein WH Department of Biochemistry and Molecular Biology University of Texas M.D. Anderson Cancer Center Box 117, 1515 Holcombe Blvd., Houston, TX 77030, USA Tel: 713 792 3646 Fax: 713 790 0329.	

```

fax: 713 750 0529.
FEATURES
  Location/Qualifiers
    1..600
      /organism="Mus musculus"
      /db_xref="taxon:10090"
      /clone_lib="Mouse E14.5 retina lambda ZAP II Library"
      /tissue_type="neural retina"
      /dev_stage="embryonic day 14.5 post-fertilization"
      /note="Vector: pAMP10 (Gibco); Cloned unidirectionally.
Primer: Oligo dT. RNA Isolation: cytoplasmic RNA preps
(Wanniatlis); Cloning Technique: CUA Cloning (CloneAmp,
Life Technologies); Average insert size: 1.8 Kb;
Insertion site: TACGCCACTGAATTCGAGTG---. Other
information regarding entire library may be found at
http://pga.swmed.edu/data/Libraries/microarray_cdna_library.
htm."
BASE COUNT      127 a  167 c  152 g  152 t      2 others
ORIGIN
Alignment Scores:
  Pred. No.:      1.38e-40      Length:      600
  Score:         401.00        Matches:      82
  Percent Similarity: 70.09%    Conservative: 0
  Best Local Similarity: 70.09%  Mismatches:   3
  Query Match:    64.16%       Indels:      33
  DB:             12          Gaps:       1

```

```

US-09-897-438B-2 (1-117) x BG6803882 (1-600)

      QY          1 GluGlnCysGlyThrIleMethHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20
      Ddb          596 GAGCAGTGCGGCACCATCATGCTGGCAATGCTGCACCTTCGTGAGCCGTACGCCCT 537
      QY          21 ArgGluLeuThrThrThrCysLeuAsnThrThrThrAlaSerValLeuGlnPheSerIle 40
      Ddb          536 CGAGAGCTGACCACCATGCTGTGAACANCAACAGCATCTGTCTCCAGCTTTTCC --- 480
      QY          41 GlySerGlySerCysArgPheSerTyrSerAspProSerIleThrValSerTyrAlaLys 60
      Ddb          480 ----- 480
      QY          61 AsnAsnThrAlaAspTrpIleGlnLeuGluLysIleArgAlaProSerAsnValSerThr 80
      Ddb          479 -----ATTGA-GCCCCCTCCAATGTGAGCAC 454
      QY          81 ValIleHisIleLeuTyrLeuProGluGluAlaLysGlyGluSerValGlnPheGlnTrp 100
      Ddb          453 GTCATCCACATCTCTAGCTCCCCGAGGAAGCAAGGGGAGAGCGTCAGTTCACGTGG 394
      QY          101 LysGlnAspSerLeuArgValGlyGluValTyrGluAlaCysTrpAlaLeu 117
      Ddb          393 AAACAGGACGCTCGAGTGGGTGAGGTGATGAGGCCTGCTGGGCCCTG 343

```


/strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-CA1-bbs-e-06-0-UI"
 /clone_lib="UI-R-CA1"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: p7T3D-fac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-CA1
 library is a subtracted library derived from the following
 tissues: thalamus, cerebellum, hypothalamus, medulla, pons
 , midbrain, cerebral cortex, corpus striatum, testis, and
 hippocampus. For a detailed description of the library
 from which this clone was derived, please visit our web
 site at rateat.eng.uiowa.edu. The subtraction has been
 previously described in (Bonaldo, Lennon and Soares,
 Genome Research 6:791-806, 1996)
 TAG_LIB=UI-R-CA1
 TAG_TISSUE=pons
 TAG_SEQ=AGCAGC"

BASE COUNT 116 a 156 c 131 g 165 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1.84e-06 Length: 568
 Score: 132.00 Matches: 23
 Percent Similarity: 96.15% Conservative: 2
 Best Local Similarity: 88.46% Mismatches: 1
 Query Match: 21.12% Indels: 0
 DB: 12 Gaps: 0

US-09-897-438B-2 (1-117) x BF387628 (1-568)

Qy 92 LysGlyGluSerValGlnPheGlnTrpLysGlnAspSerLeuArgValGlyGluValTyr 111
 ||||| ::|||
 Db 565 AAGGGAGGAGCGTCACGTCCTCAGTGGAACAGGACACCCCTCGTGCGGTGAGGTGAC 506
 Qy 112 GluAlaCysTrpAlaLeu 117
 ||||| ::|||
 Db 505 GAAGCGTCTGGGCGCTG 488

RESULT 5
 CNS012TZ
 LOCUS
 DEFINITION Tetraodon nigroviridis genome survey sequence PUC-ori end of clone
 221M15 of library G from Tetraodon nigroviridis, genomic survey
 sequence.

ACCESSION
 VERSION AL174752.1 GI:7812809
 KEYWORDS GSS: genome survey sequence.
 SOURCE Tetraodon nigroviridis.
 ORGANISM Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodon.

REFERENCE
 AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
 Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
 Saurin,W. and Weissenbach,J.
 Human gene number estimate provided by genome wide analysis using
 Tetraodon nigroviridis DNA sequence

JOURNAL
 REFERENCE Unpublished
 AUTHORS
 TITLE 2 (bases 1 to 1022)
 Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
 Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
 Weissenbach,J.
 Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis

JOURNAL
 REFERENCE Unpublished
 AUTHORS 3 (bases 1 to 1022)
 TITLE Genoscope.
 JOURNAL Direct Submission
 REFERENCE Submitted (12-APR-2000)
 COMMENT This sequence is a single read and was generated as part of a large

scale clone-end sequencing project of the Tetraodon nigroviridis
 genome. For more information, please take a look at
<http://www.genoscope.cns.fr/tetraodon>.

FEATURES
 source

Location/Qualifiers
 1..1022
 /organism="Tetraodon nigroviridis"
 /db_xref="taxon:99883"
 /clone="221M15"
 /clone_lib="G"
 /note="Genoscope sequence ID : COAG2221AG08SP1-end :
 PUC-ori"

BASE COUNT 246 a 227 c 241 g 307 t 1 others
 ORIGIN

Alignment Scores:
 Pred. No.: 0.000126 Length: 1022
 Score: 120.50 Matches: 34
 Percent Similarity: 40.82% Conservative: 6
 Best Local Similarity: 34.69% Mismatches: 10
 Query Match: 19.28% Indels: 49
 DB: 17 Gaps: 3

US-09-897-438B-2 (1-117) x CNS012TZ (1-1022)

Qy 1 GluGlnCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20
 ||||| ::|||
 Db 156 GAGCAGTGTGGGTCCTCATGCATGGAGGACCGTGACCTTCTGTGAACCCCTTGGAGAG 215
 Qy 21 ArgGluLeuThr-----ThrThrCys----- 27
 ||||| ::|||
 Db 216 CGAGAGCTGGTAAATCTATGTTGACCGTCGACTTTACATC-TGCACGGAACCCGCCA 274
 Qy 27 ----- 27
 Db 275 TCGTCACTGTGACAAAAAGTTGTTCTATGTTTGAATTTGGTTATTAGATTCTTCTG 334
 Qy 28 -----LeuAsnThrThrAlaSerValLeuGlnPheSer 39
 ||||| ::|||
 Db 335 TGTGCTGTCTAGGTACCGTGCCCTCAACACAGCAGCCTCAGTCTCAGTTTGCC 394
 Qy 40 IleGlySerGlySer-----CysArgPheSerTyrSer 50
 ::||| :::: ||| ||::|||
 Db 395 CTCGGTGAGTTTACAGATTATTTTGCAGCATTAATATGCAATCATTTTCT 448

RESULT 6
 B0942386
 LOCUS
 DEFINITION AGENCOURT_8800405 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6312366
 5', mRNA sequence.

ACCESSION
 VERSION B0942386.1 GI:22357864
 KEYWORDS EST.
 SOURCE house mouse.

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE
 1 (bases 1 to 979)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cqabps@email.nih.gov

CDNA Library Preparation: Resgen, Invitrogen Corp

CDNA Library Arrayed by: The I.M.A.G.E. Consortium

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

plate: LLAM13737 row: m column: 07

High quality sequence stop: 689.

FEATURES

source

Location/Qualifiers

1..979

[illegible]

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
 Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y.
 Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.
 e mouse tissues.

FEATURES

```

source
1. .670
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="RIKEN full-length enriched, 16 days neonate
cerebellum"
/tissue_type="cerebellum"
/dev_stage="16 days neonate"
/lab_host="DH10B"
/notes="Site_1: Sali; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGATTCGAGCTTTATTAATTAATCCCGCCCGCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I."
BASE COUNT 228 a 120 c 114 g 208 t
ORIGIN
Alignment Scores:
Pred. No.: 1.41 Length: 670
Score: 86.50 Matches: 29
Percent Similarity: 43.88% Conservative: 14
Best Local Similarity: 29.59% Mismatches: 38
Query Match: 13.84% Indels: 17
DB: 10 Gaps: 4

```

US-09-897-438B-2 (1-117) x BB628782 (1-670)

```

Qy 9 GlyAsnAlaValThrPheCysGluProTyrGlyProArgGluLeuThrThrCysLeu 28
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 282 GGAACCTTGAAGACAGCTGTGAAGCTTATCTCTGGCTCTCTGACCTCACGTGCTT 341

Qy 29 AsnThrThrAlaSerValLeuGlnPheSerlleGlySerGlySerCysArgPheSer 48
    ::: :||| ||| ||| :||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 342 CATGTAAGAGTCTCTCGGCTCTGTTGGATGTGGAAAAAGGC---TGTGATGGCAGC 398

```

```

Qy 49 TyrSer-----AspProSerIleThrValSerTyrAlaLysAsnThr 63
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 399 ATGTCACACAGACAGACAGACACTCTCTCTGAAGTCAAGCGCGTTTCTTAACAAG 458

Qy 64 AlaAspTrpIle-----GlnLeuGluLysIleArgAlaProSerAsnVal 78
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 459 CAGTGCTGGCTAATAGGATACATCTGTAGACTTTGAAAAATATTTGGCGCTATACACATCA 518

Qy 79 -----SerThrValIleHisIleLeuTyrLeuProGluGlu 90
    :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 519 TTCTGAATGAGGAAATAAAGAAATACATACATTCCTGTATCTCAAGAACA 572

RESULT 10
LOCUS B0180824 745 bp mRNA linear EST 30-APR-2002
DEFINITION UI-M-EX0-bxc-m-18-0-UI.r1 NIH_BMAP_EX0 Mus musculus cDNA clone
IMAGE:57064425 5', mRNA sequence.
ACCESSION B0180824
VERSION B0180824.1 GI:20356316
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 745)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.

```

FEATURES

```

source
1. .745
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone_image="5706425"
/clone_lib="NIH_BMAP_EX0"
/tissue_type="whole brain"
/dev_stage="embryo 15.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/notes="Organ: brain; Vector: pYX-Asc; Site_1: EcoR I;
Site_2: Not I; the library was constructed according to
Bonald, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is GTGCGTGGAA. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP); 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institutes of Mental Health
(NIMH), Hemlin Chin, Ph.D., program coordinator."
BASE COUNT 159 a 199 c 207 g 178 t
ORIGIN
Alignment Scores:
Pred. No.: 1.64 Length: 745
Score: 86.50 Matches: 32
Percent Similarity: 40.60% Conservative: 22
Best Local Similarity: 24.06% Mismatches: 52

```

```

Query Match: 13.84% Indels: 27
DB: 14 Gaps: 5
US-09-897-438B-2 (1-117) x BQ180824 (1-745)

QY 3 CysGlyThrIleMethHisGlyAsnAlaValThrPheCysGluProTyrGlyProArgGlu 22
    |||||::: ||:::||||::: || ||| |||
Db 85 TGTGGCTCCACA-----GGCGATGCCCTGGTCTTTATTGAAAGGCCAGCACCGCTTAC 138
    |||||::: ||:::||||::: || ||| |||

QY 23 LeuThrThrThrCysLeuAsnThrThrAlaSerValLeuGlnPheSerIleGlySer 42
    :::: |||||::: ||| ||||| |||
Db 139 GTGGTCACGACAGACCTCCTGTGAATGAGGACTCATCTCTACAGATAGACTTT---GCT 195
    |||||::: ||:::||||::: || ||| |||

QY 43 GlySerCysArgPheSerTyrSerAspProSerIleThrValSerTyrAlaLysAsnAsn 62
    |||||::: ||:::||||::: || ||| |||
Db 196 GCCTCCTGCTAGTCACACACTCCCTGCTATGCTATTGAATGAGGACTCGGTGATCTC 255
    |||||::: ||:::||||::: || ||| |||

QY 63 ThrAlaAspTrpIleGlnLeuGluLysIleArgAlaProSerAsnValSer---ThrVal 81
    |||||::: ||:::||||::: || ||| |||
Db 256 GGTCTGTCTGGCACCGCTGGTGAGGACTGCCTGCCTACCAATGTTGAGTGTAGTCT 315
    |||||::: ||:::||||::: || ||| |||

QY 82 IleHisIle-----SerIleGlySer-----GlySerCysArgPheSerTyr 49
    |||||::: ||:::||||::: || ||| |||
Db 316 TACCACCTGCAGGGATCCTGTGTGACAGATCTTTCAACAAGTGGACCAAGATCACTCTG 375
    |||||::: ||:::||||::: || ||| |||

QY 86 TyrLeuProGluGluAlaLysGlyGluSerValGlnPheGlnTrp----- 100
    |||||::: ||:::||||::: || ||| |||
Db 376 CCCCTGCCTTCCTACACAGGCTCAAGCCACTCGTTCCTGCTGCATCAGCCAGCGCCT 435
    |||||::: ||:::||||::: || ||| |||

QY 101 -----LysGlnAspSerLeuArgValGlyGluValTyr 111
    |||||::: ||:::||||::: || ||| |||
Db 436 TTTGACAGCAGCAGACCTGGGCAATAGATNATGCTAT 474

RESULT 11
BE892165/c 601434685F1 NIH_MGC_72 Homo sapiens cdna clone IMAGE:3919964 5',
LOCUS mRNA sequence.
DEFINITION BE892165
ACCESSION BE892165.1 GI:10352223
VERSION BE892165.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 942)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM9750 row: i column: 21
High quality sequence stop: 577.
Location/Qualifiers
1. .942
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3919964"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site:1: NotI;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 260 a 258 c 234 g 190 t

```

```

ORIGIN
Alignment Scores: 3.61 Length: 942
Pred. No.: 85.00 Matches: 33
Score: 33.33% Conservative: 19
Percent Similarity: 21.15% Mismatches: 52
Best Local Similarity: 13.60% Indels: 52
Query Match: 12 Gaps: 8
DB: 12

US-09-897-438B-2 (1-117) x BE892165 (1-942)

QY 3 CysGlyThr-----IleMethHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20
    |||||::: ||:::||||::: || ||| |||
Db 838 TGTGGCCCGCCGGCGTGTGCATAGTTCCTCCCTGGTG---CTGTCTCCGCGCGCCGCTT 782
    |||||::: ||:::||||::: || ||| |||

QY 21 ArgGluLeuThrThrThrCysLeuAsnThrThrAlaSerValLeuGlnPhe----- 38
    |||||::: ||:::||||::: || ||| |||
Db 781 CGTCTCTGTGAGGCTGCGTGTCTTTTTCGCCCTGTGCGCATTTGTTGACATACCCGTA 722
    |||||::: ||:::||||::: || ||| |||

QY 39 -----SerIleGlySer-----GlySerCysArgPheSerTyr 49
    |||||::: ||:::||||::: || ||| |||
Db 721 CGTCCGCGAGCGCTCTCTCGATTGTAATTTCTTCAGGGGATCCTGTAGACCTGTCTTA 662
    |||||::: ||:::||||::: || ||| |||

QY 50 SerAspProSerIleThrValSerTyrAlaLysAsn----- 61
    |||||::: ||:::||||::: || ||| |||
Db 661 TGAACCCCTTTAGTAACATTAGATATCTTCAACAGCATATTCTTATGGATAAGGCAC 602
    |||||::: ||:::||||::: || ||| |||

QY 62 -----AsnThrAla 64
    |||||::: ||:::||||::: || ||| |||
Db 601 AGCTTCCAGGATGCTGCGCTCTGGATGTTTCAGAGGAGGAGAAAGGTGTCAAAACAGCG 542
    |||||::: ||:::||||::: || ||| |||

QY 65 AspTrpIleGlnLeuGluLys-----IleArgAlaProSerAsnValSerThr 80
    |||||::: ||:::||||::: || ||| |||
Db 541 GACTCGTTAAATTTTCCCAAAACTGTTGTTGTCGAGGACCAAGCA----- 494
    |||||::: ||:::||||::: || ||| |||

QY 81 ValIleHisIleLeuTyrLeuProGluAlaLysGlyGluSerValGlnPheGlnTrp 100
    :::: |||||::: ||| ||| |||
Db 493 -----TGGCAGCTGTCCTCCCATCTGCCACAGAGCATGCAAAAGGCAATGG 446
    |||||::: ||:::||||::: || ||| |||

QY 101 LysGlnAspSerLeuArgValGlyGluValTyrGluAla---CysTrp 115
    :::: |||||::: ||| ||| |||
Db 445 CGGCACGATGACAGAGGGGTGGCCACCAAGTATCATCTGTCTGTTGG 398
    |||||::: ||:::||||::: || ||| |||

RESULT 12
AW158822 LOCUS 369 bp mRNA linear EST 05-NOV-1999
DEFINITION za45g06.x1 Xenopus EST library Xenopus laevis cdna clone za45g06
5', mRNA sequence.
ACCESSION AW158822
VERSION AW158822.1 GI:6270851
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 369)
AUTHORS Schutz,K., de la Bastide,M., Huang,E.N., Nascimento,L., Preston,R.,
Shah,R., Swaby,I., Shekher,M., Spiegel,L., Vil,M.D. and McCombie
,W.R.
Expressed sequence tags from Xenopus
Unpublished (1999)
JOURNAL
COMMENT Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874
Email: mcombie@cshl.org
Plate: za45 row: g column: 06
Seq primer: M13 universal forward primer
High quality sequence stop: 369.

```

FEATURES

source	Location/Qualifiers	
	i..369	/organism="Xenopus laevis"
		/db_xref="taxon:8355"
		/clone="za45g06"
		/clone_lib="Xenopus EST library"
		/tissue_type="total brain tissue"
		/cell_line="W22-RGA"
		/dev_stage="tadpole"
		/notes="vector: Lambda zap I; Site_1: XbaI; This library was supplied by Holly Cline (Cold Spring Harbor Labs). cDNA synthesis with oligo dT Xba I (Xba I cloning site) RNA: stage 50-56 tadpoles, total brain tissue, GFC extraction method."

BASE COUNT	76 a	69 c	92 g	132 t
ORIGIN				

Alignment Scores:		Pred. No.:	Length:	Score:
			Matches:	369
			Conservative:	34
			Mismatches:	16
			Indels:	51
			Gaps:	26
				7

US-09-897-438B-2 (1-117) x AW158822 (1-369)

QY	3	CysGlyThrIleMethHisGlyAsnAlaValThrPhe-----CysGluProTyrGly	19
Db	15	TGTGGTGCTGGCGCTCTGGATGGCGCTTCAC TTCAGTGGGGATGCAGT-----	65
QY	20	ProArgGluLeuThrThrCysLeuAsnThrThrAlaSerValLeuGlnPheSer	39
Db	66	---CGTTTGGTTCTGTGGATCTCAATCTCAATGCTGAGTTGTTTCAGTTTTAT	122
QY	40	IleGlySerGlySerCysArgPheSerTyrSer---AspProSerIleThrValSerTyr	58
Db	123	TTCATGTACGGCTGCCCTCATAAACACCTCACAGCCGTAATCAAGGAGTCTCTCGAGTAT	182
QY	59	AlaLysAsnAsnThrAlaAspTrpIleGlnLeuGluLysIle-----Arg	73
Db	183	TC TGTGTAATGGGGTATTCAATTGTTCTTTGTATGAAAATCTTTTACGATCAGTACAGC	242
QY	74	AlaProSerAsnValserThrValIleHisIleLeuTyrLeuProGluGluAlaLysGly	93
Db	243	AAACCTGGTTTGTGAAT-----ATTCTTTTGCTGCTAGTGCCAAGACA	287
QY	94	GluSerValGlnPheGlnTrp-----LysGlnAspSerLeuArgValGlyGluVal	110
Db	288	GTGGGAACCGGTTTCCGCTGTGGCAACCAAAACATGATGGCTGGACC AAAATGAT---	344
QY	111	TyrGluAlaCysTrpAlaLeu	117
Db	345	-----TGGGCAATT	353

RESULT	13	B1739930	674 bp	mRNA	linear	EST 20-SEP-2001
LOCUS	B1739930	603362473F1 NIH_MGC_94	Mus musculus	cDNA clone	IMAGE:5369371	5'
DEFINITION	mRNA sequence.					
ACCESSION	B1739930					
VERSION	B1739930.1 GI:15716943					
KEYWORDS	EST.					
SOURCE	house mouse.					
ORGANISM	Mus musculus					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus					
AUTHORS	1 (bases 1 to 674)					
TITLE	NIH-MGC http://mgc.nci.nih.gov/					
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)					
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabps-re@mail.nih.gov					

Tissue Procurement: The Cepko Laboratory					
cDNA Library Preparation: Life Technologies, Inc.					
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)					
DNA Sequencing by: Incyte Genomics, Inc.					
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM1940 row: a column: 20 High quality sequence stop: 635. Location/Qualifiers					
FEATURES					
source					
1..674					
/organism="Mus musculus"					
/db_xref="taxon:10090"					
/clone="IMAGE:5369371"					
/clone_lib="NIH_MGC_94"					
/tissue_type="retina"					
/lab_host="DH10B (phage-resistant)"					
/note="Organ: eye; Vector: pCMV-SPORT6; Site: NotI; Site2: SalI; Cloned unidirectionally; oligo-dT primed Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies Note: this is a NIH_MGC Library."					
BASE COUNT 150 a 177 c 189 g 158 t					
ORIGIN					
Alignment Scores:					
Pred. No.: 2.94 Length: 674					
Score: 84.00 Matches: 34					
Percent Similarity: 43.55% Conservative: 20					
Best Local Similarity: 27.42% Mismatches: 50					
Query Match: 13.44% Indels: 20					
DB: 13 Gaps: 8					
US-09-897-438B-2 (1-117) x BI739930 (1-674)					
Qy	3	CysGlyThrIleMethHisGlyAsnAlaValThrPhe-----CysGluProTyGly	19		
Db	194	TGTGGGCTGTGGCTTCGGCTCGGCTCTCCATTCAAGTGGGGCTGCAGC-----	244		
Qy	20	ProArgGluLeuThrThrThrCysLeuAsnThrThrThrAlaSerValLeuGlnPheSer	39		
Db	245	--CGATGTGTACATCTGGATCTGAACCTCACCAATGCTGAGTTTATCCAGTTTAC	301		
Qy	40	IleGlySerGlySerCysArgPheSerTySer-----AspProSerIleThrValSer	57		
Db	302	TTATGTATGGA---TGCCCTCATTCACCCGCCAACCGTAACAGGGAGTCTCTCTGGAG	358		
Qy	58	TyrAlaLysAsnAsnThrAlaAspTrp---IleGlnLeuGluLysIleArgAlaProSer	76		
Db	359	TACTCTGCAATGGAGCATCACCTGGAACTGCTGATGAGATGTTCTATGACCAGTAC	418		
Qy	77	AsnValSerThrValIleHisIleLeuTyLeuProGluGluAlaLysGlyGluSerVal	96		
Db	419	AGCAAACCTGGATTGTGAATATCCCTT---CTCCCCTCATGCTAAAGAGATGCCACT	475		
Qy	97	GlnPheGlnTrp-----LysGlnAspSerLeuArgValGlyGluValTyrGluAla	113		
Db	476	CGCTTCGATGGTGGCAGCACGACATGATGGCCTTGACCAGATGAC-----	523		
Qy	114	CysTrpAlaLeu 117			
Db	524	--TGGGCCATT 532			
RESULT 14					
BQ769321					
LOCUS					
DEFINITION					
IMAGE:5721422 5', mRNA sequence.					
ACCESSION					
VERSION BQ769321					
KEYWORDS EST.					
SOURCE					
ORGANISM					
house mouse					
Mus musculus					
692 bp mrna linear EST 26-JUL-20					
UT-M-FIO-byp-n-15-0-UI.r1 NIH_BMAP_F10 Mus musculus cDNA clone					

Db	476	AGCAACCTGGATTGTGAAT			---	ATCCTTCTCCTCTCTGATGCTAAA	520
Qy	93	GlyGluSerValGlnPheGlnTrp	-----	-----	-----	LysGlnAspSerLeuArgValGlyGlu	109
Db	521	GAGATGCCACTCGTTCGATGGCAGCCAGCATGATGGCTTGACCAAGATGAC			---	---	580
Qy	110	ValTyrGluAlaCysTrpAlaLeu	117		---	---	---
Db	581	-----	-----	-----	-----	TGGGCCATT	589
RESULT	15						
BO771419							
LOCUS	BO771419	706 bp	mrna	linear	EST	26-JUL-	
DEFINITION	UT-M-F10-byu-c-12-0-UL.r1 NIH-BMAP_F10 Mus musculus cDNA clone IMAGE:5702339 5', mRNA sequence.						

VERSION	BQ7/1419.1	GI:21979895
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
TITLE	1 (bases 1 to 706)	
JOURNAL	NH-MGC http://mgc.nci.nih.gov/ .	
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: Dr. Jim Lin, University of Iowa cDNA Library preparation: Dr. M. Bento Soares, University of Iowa cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov This clone was contributed by the Brain Molecular Anatomy Project (BMAP)	

```

/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5702339"
/clone_lib="NIH_BMAP_F10"
/tissue_type="whole brain"
/dev_stage="embryo 12.5dpc"
/lab_host="DHI10B (T1 phage resistant)"
/site_1="EcoR I;
/site_2="Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CAGCACGAC. This library was created for the University
of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."
BASE COUNT      153 a   190 c   190 g   173 t
ORIGIN

Alignment Scores:
Pred. No.:      3.15      Length:      706
Score:          84.00      Matches:    34
Percent Similarity: 41.41%      Conservative: 19
Best Local Similarity: 26.56%      Mismatches: 47
Query Match:    13.44%      Indels:    28

```



```

DB:          14          Gaps:          8
US-09-897-438B-2 (1-117) x BQ771419 (1-706)
Qy   3 CysGlyThrIleMetHisGlyAsnAlaValThrPhe-----CysGluProTyrGly 19
      ||||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db  311 TGTGGGCTGTGGCTTCGGCTGGCTCTCCATTTCAGTGGGGGCTGCAGC----- 361
Qy   20 ProArgGluLeuThrThrCysLeuAsnThrThrAlaSerValLeuGlnPheSer 39
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  362 ---CGATTGTTAGTCACGTGTGGATCTGAACCTCACCAATGCTGAGTTTATCCAGTTTAC 418
Qy   40 IleGlySerGlySerCysArgPheSerTyrSer-----AspProSerIleThrValSer 57
      ||| ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db  419 TTTATGTATGGA---TGCCTCATTACGCCGAGCAACCGTAACACGGAGTCTGCTGGAG 475
Qy   58 TyrAlaLysAsnAsnThrAlaAspTrp-----IleGlnLeuGluLysIle 72
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db  476 TACTCTGTCAATGGAGGCATCACCTGGAACCTTGCTGATGGAGATTTCTATGACCACTAC 535
Qy   73 ArgAlaProSerAsnValSerThrValIleHisIleLeuTyrLeuProGluGluAlaLys 92
      ||| ||| ||||| ||||| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db  536 AGCAACCTGGATTGTGAAT-----ATCCTTCTCCCTCTCGATGCTATAA 580
Qy   93 GlyGluSerValGlnPheGlnTrp-----LysGlnAspSerLeuArgValGlyGlu 109
      :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db  581 GAGATTGCCACTCGCTTCCGATGTGGCAGCCACGACATGATGGCCTTGACCAGAAATGAC 640
Qy  110 ValTyrGluAlaCysTrpAlaLeu 117
      ||||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db  641 -----TGGGCCATT 649

```

Search completed: November 6, 2002, 20:53:32
Job time : 2222 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 6, 2002, 13:07:41 ; Search time 2762 Seconds
(without alignments)
3698.438 Million cell updates/sec

Title: US-09-897-438B-1
Perfect score: 351
Sequence: 1 gagcagtgtggcaccatcat.....atgaggcctgtgggccttg 351

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0
Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	351	100.0	11673	6	AX3053377	AX3053377 Sequence
2	351	100.0	11673	10	MMU24703	U24703 Mus musculus
3	299.8	85.4	11187	10	AB049473	AB049473 Rattus no
4	263.6	75.1	11580	6	AX410790	AX410790 Sequence
5	263.6	75.1	11580	9	HSU79716	U79716 Human reeli
6	263.6	75.1	11580	11	G30936	G30936 SWS2926 Er
7	263.6	75.1	11580	11	G30938	G30938 SWS3176 Er
8	189.4	54.0	10634	5	AF090441	AF090441 Gallus ga
9	135	38.5	163985	10	AC121878	AC121878 Mus muscu
10	135	38.5	163985	2	AC023062	AC023062 Mus muscu
11	115.8	33.0	126130	2	AC095877	AC095877 Rattus no
12	115.8	33.0	202764	2	AC128022	AC128022 Rattus no
13	114.2	32.5	183641	2	AC124933	AC124933 Rattus no
14	98.8	28.1	93163	9	HSAC000121	AC000121 Human BAC
15	53.8	15.3	185996	2	AC041023	AC041023 Homo sapi
16	52.2	14.9	158	4	AF232904	AF232904 Bos tauru
17	37.2	10.6	2662	5	GGA317960	AF17960 Gallus ga
18	37.2	10.6	145264	9	AC107939	AC107939 Homo sapi
19	36.6	10.4	90166	2	AC114052	AC114052 Rattus no
20	36.6	10.4	143517	9	AL356752	AL356752 Human DNA
21	36.6	10.4	195670	2	AC115449	AC115449 Rattus no
22	36.4	10.4	55452	3	AC004328	AC004328 Drosophil
23	36.4	10.4	63751	2	AC017489	AC017489 Drosophil
24	36.4	10.4	164567	3	AC091500	AC091500 Drosophil
25	36.4	10.4	246095	3	AE003802	AE003802 Drosophil
26	35.8	10.2	173841	10	AL732391	AL732391 Mouse DNA
27	35.8	10.2	189541	10	AC093447	AC093447 Mus muscu
28	35.6	10.1	1365	9	HSJ308525	HSJ308525 Homo sapi
29	35.6	10.1	1812	10	AF245444	AF245444 Mus muscu
30	35.6	10.1	2619	9	HSJ299451	HSJ299451 Homo sapi
31	35.6	10.1	2973	9	HSJ249210	HSJ249210 Homo sapi
32	35.6	10.1	3200	10	RATGLUR7A	M83552 Rattus norv
33	35.6	10.1	3344	6	AR001481	AR001481 Sequence
34	35.6	10.1	3380	11	G26866	G26866 human STS S
35	35.6	10.1	3593	10	RNGLUR7	Z11716 R. norvegicu
36	35.6	10.1	3614	9	HSU16127	U16127 Human gluta
37	35.6	10.1	3619	6	AR180924	AR180924 Sequence
38	35.6	10.1	3633	10	AF027331	AF027331 Rattus no
39	35.2	10.0	14089	10	AF060195	AF060195 Mus muscu
40	35.2	10.0	34849	10	AB053120	AB053120 Mus muscu
41	35.2	10.0	148846	10	AL591204	AL591204 Mouse DNA
42	35	10.0	87370	2	AL356412	AL356412 Homo sapi
43	35	10.0	153056	10	AL607090	AL607090 Mouse DNA
44	35	10.0	160557	9	AC117945	AC117945 Homo sapi
45	35	10.0	193468	2	AC125591	AC125591 Rattus no

ALIGNMENTS

RESULT 1
AX3053377
LOCUS AX3053377 11673 bp DNA linear PAT 11-DEC-2001
DEFINITION Sequence 128 from Patent WO0188188.
ACCESSION AX3053377
VERSION AX3053377.1 GI:17644926
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1
AUTHORS Ishikawa, K., Asai, S., Takahashi, Y., Nagata, T. and Ishii, Y.
TITLE Method for examining ischemic conditions
JOURNAL Patent: WO 0188188-A 128 22-NOV-2001;

School Juridical Person Nihon University (JP)

FEATURES
source
1. .11673
/organism="Mus musculus"
/db_xref="taxon:10090"

BASE COUNT 2831 a 2985 c 2985 g 2872 t

5'UTR
CDS

Query Match 100.0%; Score 351; DB 6; Length 11673;
Best Local Similarity 100.0%; Pred. No. 5.7e-96;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAGCAGTGGCAGCATCATGCGAATGCTGTCACCTTCTGTGAGCCGTACGGCCCT 60
Db 970 GAGCAGTGGCAGCATCATGCGAATGCTGTCACCTTCTGTGAGCCGTACGGCCCT 1029
Qy 61 CGAGAGCTGACACACATGCTGAACACACACACATGCTGCTCCAGTTTCCATT 120
Db 1030 CGAGAGCTGACACACATGCTGAACACACACACATGCTGCTCCAGTTTCCATT 1089
Qy 121 GGGTCAGGATCATGTCGATTAGTTACTTCTGACCCAGCATCACTGTGTCATAGCCCAAG 180
Db 1090 GGGTCAGGATCATGTCGATTAGTTACTTCTGACCCAGCATCACTGTGTCATAGCCCAAG 1149
Qy 181 AACATACCGTGATTGATTCAGCTGGAGAAATAGAGCCCTTCCAATGTGAGCACA 240
Db 1150 AACATACCGTGATTGATTCAGCTGGAGAAATAGAGCCCTTCCAATGTGAGCACA 1209
Qy 241 GTCATCCACATCTGTACTCTCCGAGGAGCAAGGAGAGCGTGCAGTCCAGTGG 300
Db 1210 GTCATCCACATCTGTACTCTCCGAGGAGCAAGGAGAGCGTGCAGTCCAGTGG 1269
Qy 301 AAACAGGACACCTCGAGTGGGTGAGTGATGAGCCCTGCTGGGCCCTG 351
Db 1270 AAACAGGACACCTCGAGTGGGTGAGTGATGAGCCCTGCTGGGCCCTG 1320

RESULT 2
MMU24703 11673 bp mRNA linear ROD 19-DEC-1997
DEFINITION Mus musculus reelin mRNA, complete cds.
ACCESSION U24703
VERSION U24703.1 GI:2702252

KEYWORDS
SOURCE
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 11673)
D'Arcangelo, G., Miao, G.G., Chen, S.C., Soares, H.D., Morgan, J.I. and Curran, T.
A protein related to extracellular matrix proteins deleted in the mouse mutant reeler.
Nature 374 (6524), 719-723 (1995)
95231649
7715726

REFERENCE 2 (bases 1 to 11673)
D'Arcangelo, G.
Direct Submission
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

REFERENCE 3 (bases 1 to 11673)
Submitted (10-APR-1995) Roche Institute of Molecular Biology, 340 Kingsland St., Nutley, NJ 07110, USA
D'Arcangelo, G.
Direct Submission
AUTHORS
TITLE
JOURNAL

REFERENCE 4 (19-DEC-1997) Dev. Neurobiology, St. Jude Childr. Res. Hosp., 332 N. Lauderdale, Memphis, TN 38105, USA
Submitted by submitter
Sequence update by submitter
On Dec 19, 1997 this sequence version replaced gi:902486.
COMMENT
FEATURES

source
1. .11673
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="5"

3'UTR
polyA_signal
BASE COUNT 2831 a 2985 c 2985 g 2872 t
ORIGIN

/map="D5Gmr1"
/tissue_type="cerebellum"
/dev_stage="adult"
1. .282
283. .10668
/codon_start=1
/product="reelin"
/protein_id="AAB91599.1"
/db_xref="GI:2702253"
/translation="MERGQWAPRALVAVLLIATLARAATGYPRFSPFFFLCTHH
GELEGGEGEVLSLHAGNPYYVPGQYHVTIISTFFDGLLVLTSTISQSS
QSIGGSAFAGIMSDHGFQCMCVSAVSHVSLPTNLNLFVAPPAGTGCVMFMA
TATHRGQVIFKDALAQCLCEQGAETATYSHLAEHSDSVILRDFDPSYQLEAFN
IWECSNCEMGECGTTIMHGNATFCEPRLTTLTCLNTTASLQFSIGSGSCRF
SYDPSITVYAKNTADMIOLKIRAPSNVSTVILYLPBEAKESQVQWOKDLSL
RGEVVEACWALDNILVINSABEVLEONLDPDGNMLFFPGATVKKSCSDGNSI
YFNGEGSEFNATTROVDLSTEDIEQWSEFESEPQTGWDILGAVVAGDCGTESGL
SLVFLDKGERKKLTGYMNTGYNLRFYFVMGGICDPGVSHENDIILYAIKGRKEHI
ALDTLYSSYKVPKSLVSVINPELQTPATKFNKRSKSHOQYNNVAVDFHVLPLYLP
STMSHMIQFSINLGGCTHQPNSVLEFSTNKRKWSLLHTECLPEICAGPLHPSTV
YSSNTSGWNRITIPLPNAALTDRTRIWRQTGPILGNMAIDNVYISGAEKFCGSR
GOTRHKCKDCPGFCGACEMASQTPMFISESFGSARLSSVHNYSIRGBAUSFCGG
VLASGKALVFNKDRRLQTLTSLDSSQSRFLQTLRLGSKSVLSTCRADPQEGEVLL
HYSYDNGITWKLLEHYSYVNHPIISVELPDDARQFQIFRWMPYHSSQGEDWMA
IDIVMTSVLNFSLIDFTNLVETVSLGFLGNVQPCGHDTLCTGDSKLASSMR
YVETQSMQIGASYMIQFSLVMGQKTYTHMDNQVKLEYSAHSHYTLHWLQEBCLPSM
PSCQETASISYHASEFTOWRVTVYLPKOTWSGATFRWSQSYTYAQDEWALDNIYI
GOCPNMCSHGSCDHGCRDOGYGCTHCPAALPSTIMSDFENPSSWESQOEVI
GEBVPEOGCGVVSQSSLYFSKAGKRLQVLDLSTSWDFVQFYLIQIGESAACNK
PSREGEILQYNNNGGIQWHLAEMFSDFSKPREVYIELPAAGTPTCFRFRWPKPV
FSGEDYQWAVDDIILSEKQOVIPVNPRTLQNFYKPAEDPKNQMSVWMLANE
GMAKNDSCATTSPAMVFGKSDGDFRATRLTLKPYVLQKLNIGCTSQFSTAPV
LIQYSHDAGMSWFLKEGCFPASAAGCEGNSRELSEPTVYVYTGDEEWTITAIAPR
SLASKTRFRWIFQESSOKNVPFGLDGVYISPCPSYCGSGHCDLISGVCFCDLGYTA
AOGTCVSNTPNHSEMEDREGKLSPLWYKITGGQVCTGGTLDNGRSILFENGLKREA
RTVPLTRNISLVQFYIQTGKSTGYITPRARYBGLVYVQSDNGNGLIHLHLEUDF
MSFLEPIIISIDLPREAKTATAFRMOQHGKSAOWALGDVLIGNDSSQTFQDOK
LDGSDIQANWYRIGQGVQIDICLSMDTALITENIGNPRYAEFTWDFHVSSEFLQWE
MNGCSKPFSGAHQIOQLSNNGKDWLVTECVPTTICVHYTETSYTSRFFQNW
RVTVYPLATNSPRTFRWITNTYVGADSWAIDNVILASGCPWMSGRCIGDSCGR
VCDRGGPGGCVPVVPLPSILKDDFNGLHPDLPEVYGAERGLNGETIKSGTCLIF
KBEGLMLISRLDCTNTWTVQSLRFIAKGTTPERSHLSILQFVSQGVVTHLWDEFL
FPTQTSILFINVPLPYGAQTNATRFRLMOPYNNGKKEIWIIDDFIIGNNLLNPVLL
LDTDFEPREDNFFYPGGNIGLYCPSKGAPEEDSAMVFSVEGHSITTRDLSV
NENTIIQFEINVCSDSSADPVRLEFSDGFATWHLLPLCYHSSSLVSLCSTEHE
HPSYVYAGTQGWREVYFGLKHLGCVFRFRWYOGFYPCAGOPVTWADNVYIGQP
CEEMCYGHGSCINGKICIDPGYSGTCKISTKNPDLKDDPEGQLESREFLLMSGGK
PBRKCGILSGNNLFFNEDGLRMLVTRDLSDSHAREVQFFMRIGCCGKQVDPDRSQVPL
LQYSLNGLSWSLLQEFLEFSNNSVGRYIALEMLKARSGSTRLRMWQPSENGHYSP
WYIDQILIGNISGNTLVDDFSTLDSRWKLLHPGGTKMPVCGSTGSDALVLEKASTR
YVVTDDIAVNEDSFLQIDFAACSVDTSYATILEYYSVDLGLSWHPLVRDCLPTNVEC
SRYLQRIILVSDTFNKNWTRITLPLPSYTRSOATRFRWHPAPDFKOOTWADNVYIGD
GCLDMSGHRCVQCSQVDEQWGLLYCDEPETSILPTQKDNFNRAQSNQMLVTSVGG
KLSTVCGAVASGLAHLHGGCRLVTVDLNLTNAEYFIQFYFMGLIPLTSPNRNQVL
LEYSYNGITWNLMLFIYDYYSKPGFNILLPPDAKEATFRAFRMWQPRHGLQDNW
AIDNVILIGSADQRTVMDTFSSAPVQPHQEPADAPGVGRIATFEMFLBDKTSVNEWM
LFHDDCTVERFCDSPDGMCLGSHDGRVYAVTHDLTPENMLMOKFISVGCCKVPEKI
AQONTHQVQSTDFGYSWYLVQCLPADPKCSGSYSQSPVPEFTGKWKRTITPLPSL
TGNVRFRTYQKTSQVQWADNFIYLPGLCDNCGGGDCLEKQECIDPGYSGPNCTYL
HSLKTLKRFEDSEIEKPLDMLWSLGGSTCEGVALENTALYFGSTYRQAITQDLD
LRGAKFLQYWRIGRISNNMTSCHRPVRCGEVLLDFTSDGTITWTLHMDQKTLISV
RHDYTLPEGALNTFRLRWMPFVINSGLVSVGRVQAWALDNILIGAEINPSOLV
DTFDDGSSHEENWSPYNAVRTAGFCNPSPFLYWPKNKKDKKTINALSSRELITQPG
YMKQPIVYCCATSCGDLHSHVMEYTKDARSQVQLVQVQCLPSSNSIGCSPPQFH
EATIIYNAVSSNKKRITIQLPDHSVSSATQFRWIOKEETEKEOSWAIDHYIGEACP
LCSGHYCTTGAVCTDESDQDCSVFSEHSPYIKDNFSAVYTAHWETIQQGCPK
GSCGGLAPYAHGDSLYNGCQIRQAATKPLDTRASKIMFYLGTSQPSQSCNSDL
SGCPHTVDRKAVLLQYSVNNGITTWHVTAHQPKDFOAORVSNVPLKRMKGLVLRMWQ
PRHNGTGHQWALDHVEVVLVSTRKQNTYMNWNSRQHLRHFYNNRRRSRLRPY"

Query Match 100.0%; Score 351; DB 10; Length 11673;
Best Local Similarity 100.0%; Pred. No. 5.7e-96;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCAGTGTGGCACCACCATCATGCGTCAATGCTGTGTCACCTTCTGTGAGCGCCTACGGCCCT 60
|||||
Db 970 GAGCAGTGTGGCACCACCATCATGCGTCAATGCTGTGTCACCTTCTGTGAGCGCCTACGGCCCT 1029
|||||

QY 61 CGAGAGCTGACACACACATGCTGAAACACAAACAGCAGTCTGTCTCTCCAGTTTCCATT 120
|||||
Db 1030 CGAGAGCTGACACACACATGCTGAAACACAAACAGCAGTCTGTCTCTCCAGTTTCCATT 1089
|||||

QY 121 GGCTCAGGATCATGCTGATTTAGTCTACTCTGACCCAGCATCATGCTGTGATACGCGCAAG 180
|||||
Db 1090 GGCTCAGGATCATGCTGATTTAGTCTACTCTGACCCAGCATCATGCTGTGATACGCGCAAG 1149
|||||

QY 181 AACAAATACCGCTGATTGATTCAGCTGGAGAAAATTAGAGCCCTTCCCAATGTGAGCACA 240
|||||
Db 1150 AACAAATACCGCTGATTGATTCAGCTGGAGAAAATTAGAGCCCTTCCCAATGTGAGCACA 1209
|||||

QY 241 GTATCCACATCTCTGACCTCCCGGAGAAACGAGGAGAGCGTGCAGTTCCAGTGG 300
|||||
Db 1210 GTATCCACATCTCTGACCTCCCGGAGAAACGAGGAGAGCGTGCAGTTCCAGTGG 1269
|||||

QY 301 AAACAGACAGCCTCGCAGTGGGTGAGGTGATGAGCGCTGCTGGGCCCTG 351
|||||
Db 1270 AAACAGACAGCCTCGCAGTGGGTGAGGTGATGAGCGCTGCTGGGCCCTG 1320
|||||

RESULT 3
AB049473
LOCUS AB049473 11187 bp mRNA linear ROD 01-DEC-2001
DEFINITION Rattus norvegicus mRNA for reelin, complete cds.
ACCESSION AB049473
VERSION AB049473.1 GI:17221617
KEYWORDS
SOURCE Rattus norvegicus juvenile cerebellum cDNA to mRNA.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1
Kikkawa, S. and Terashima, T.
Rat reelin (Rein) complete CDS
Published Only in Database (2001)
2 (bases 1 to 11187)
Kikkawa, S. and Terashima, T.
Direct Submission
Submitted (29-SEP-2000) Satoshi Kikkawa, Kobe University School of
Medicine, Department of Anatomy; Chuo-ku, Kusunoki-cho 7-5-1, Kobe,
Hyogo 650-0017, Japan (E-mail: skikkawa@med.kobe-u.ac.jp,
Tel:81-78-382-5325, Fax:81-78-382-5328)

FEATURES

source
1. .11187
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/tissue_type="cerebellum"
/dev_stage="juvenile"
1. .11187
/gene="Rein"
341..10729
/gene="Rein"
/codon_start=1
/product="reelin"
/protein_id="BAB78470.1"
/db_xref="GI:17221618"
/translation="MERCWAPRTLVALLVLLLLLRARATGYPRSPFFFLCETH
HGELEGDEQGEVLISLHAGNPTIYVQGEYHVTISTFDGLLVLTGLYSTSIQS
SQSGGSGAGGIMSDHQFNQFVSVASHVSHLPTNLFSVMFIAPAGGCVNFM
ATATRGVQFKDALAQOLCEOGATEATAYSHLAEIHSDSVILRDDPSYHOLELNP
N1WAECSNCDTCEOGCTLMHGNVTFCEPYGPRELTYLNTTASVLQFSIGSGSCR
FYSYDPSIIVSYAKNNTADWIOLEKIRAPSNVSTIIHILYLPEDAKGENVQFQWKQDS

gene

CDS

QY 1 GAGCAGTGTGGCACCACCATCATGCGTCAATGCTGTGTCACCTTCTGTGAGCGCCTACGGCCCT 60
|||
Db 1031 GAACAGTGTGGCACCACCATCATGCGTCAATGCTGTGTCACCTTCTGTGAGCGCCTACGGCCCT 1090
|||

QY 61 CGAGAGCTGACACACACATGCTGAAACACAAACAGCAGTCTGTCTCTCCAGTTTCCATT 120
|||||

Db 1091 CGAGAGCTGACACACACATGCTGAAACACAAACAGCAGTCTGTCTCTCCAGTTTCCATT 1150
|||||

QY 121 GGCTCAGGATCATGCTGATTTAGTCTACTCTGACCCAGCATCATGCTGTGATACGCGCAAG 180
|||||

Db 1151 GGCTCAGGATCATGCTGATTTAGTCTACTCTGACCCAGCATCATGCTGTGATACGCGCAAG 1210
|||||

BASE COUNT
ORIGIN

Query Match 85.4%; Score 299.8; DB 10; Length 11187;
Best Local Similarity 90.9%; Pred. No. 2.5e-80;
Matches 319; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY	181	AAACAATACCGCTGATTGGATTACGCTGGAGAAAATTAGACCCCTTCCAAATGTGAGCAC	240
Db	1211	AAACAATACCGCTGATTGGATTACGCTGGAGAAAATTAGACCCCTTCCAAATGTGAGCAC	1270
QY	241	GTCAATCACATCTGTACCTCCCGAGGAGCCAAAGGGGAGAGCGTGCAGTTCACGTGG	300
Db	1271	ATCAATCACATCTGTACCTCCCGAGGAGCCAAAGGGGAGAGCGTGCAGTTCACGTGG	1330
QY	301	AAACAGGACAGCCTGCCAGTGGGTGAGGTGATGAGGCCCTGCCTGGGCCCTG	351
Db	1331	AAACAGGACAGCCTGCCAGTGGGTGAGGTGATGAGGCCCTGCCTGGGCCCTG	1381
RESULT 4			
AX410790		11580 bp	DNA
LOCUS	AX410790	Sequence 3437 from Patent WO0229103.	linear
DEFINITION	AX410790	PAT 14-JUN-2002	
ACCESSION	AX410790		
VERSION	AX410790.1	GI:21443495	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.		
JOURNAL	Gene expression profiles in liver cancer		
Patent:	WO 0229103-A 3437 11-APR-2002;		
GENE LOGIC	INC (US)		
FEATURES	Location/Qualifiers		
Source	1. .11580		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/note="ENBL/GenBank Accession No. 079716"		
BASE COUNT	3014 a 2696 c 2753 g 3116 t		1 others
ORIGIN			
Query Match	75.1%;	Score 263.6;	DB 6;
Best Local Similarity	84.6%;	Pred. No. 3e-69;	
Matches	296;	Conservative 0;	Mismatches 54;
		Indels 0;	Gaps 0;
QY	1	GACAGTGTGGCACCATCATGCGTACGCTGACCTTCTGTGAGCGGTACGGCCCT	60
Db	860	GAACAGTGTGGCGGATATGATGCGCAATGCGCTGACCTTCTGTGAACCATATGCCCCA	919
QY	61	CGAGAGCTGACCAACACATCCCTGAACACAAACAGCATCTCTCTCCAGTTTTCATT	120
Db	920	CGAAGCTGATACACAGCCCTTAATACACACAGCTTCTCTCTCCAAATTTTCATT	979
QY	121	GGTCAAGGATCATGTCGATTAGTTACTCTGACCCAGCATCATCTGTGCATACGCCAAG	180
Db	980	GGTCAAGTTCATGTCGCTTTAGTTATTTCAGACCCAGCATCATCTGTATATGCCAAG	1039
QY	181	AAACAATACCGCTGATTGGATTACGCTGGAGAAAATTAGAGCCCTTCCAAATGTGAGCAC	240
Db	1040	AATAACTCTGCGACATGGATTAGCTAGAGAAAATTAGAGCCCTTCCAAATGTGAGCAC	1099
QY	241	GTCAATCACATCTGTACCTCCCGAGGAGCCAAAGGGGAGAGCGTGCAGTTCACGTGG	300
Db	1100	ATCATCATATCTCTACCTTCTCTGAGGACGCCAAAGGGGAGAGATGTCCAATTTCACTGG	1159
QY	301	AAACAGGACAGCCTGCCAGTGGGTGAGGTGATGAGGCCCTGCCTGGGCCCTG	350
Db	1160	AAGCAGGAAAATCTTCGTGTAGTGAAGTGTATGAAGCCTGCTGGGCCCTT	1209
RESULT 5			
HSU79716		11580 bp	mRNA
LOCUS	HSU79716	Human reelin (RELN)	mRNA, complete cds.
DEFINITION	U79716		
ACCESSION	U79716		
VERSION	U79716.1	GI:1743884	
KEYWORDS			

SOURCE	Homo sapiens.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE	1 (bases 1 to 11580)
JOURNAL	DeSilva,U., D'Arcangelo,G., Braden,V.V., Chen,J., Miao,G.G.,
MEDLINE	Curran,T. and Green,E.D.
PUBMED	The human reelin gene: Isolation, sequencing, and mapping on
AUTHORS	chromosome 7
TITLE	Genome Res. 7 (2), 157-164 (1997)
JOURNAL	97202106
MEDLINE	9049633
PUBMED	2 (bases 1 to 11580)
AUTHORS	Curran,T. and Green,E.D.
TITLE	Direct Submission
JOURNAL	Submitted (26-NOV-1996) National Center for Human Genome Research,
FEATURES	National Institutes of Health, 49 Convent Drive, MSC4431, Bethesda,
Source	MD 20892, USA
	Location/Qualifiers
	1. .11580
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="7"
	1. .11580
	/gene="RELN"
	176. .10558
	/gene="RELN"
	/codon_start=1
	/product="reelin"
	/protein_id="AAC51105.1"
	/db_xref="GI:1743885"
	/translation="MERSGAAQTFLALLGLATLRARAAAGYVPRFSPFFFLCTHHG
	ELEGDEGEVLIHLIAGNPVYVPGQYHVTISTSTPFDGLLTVGLTYSVQASQ
	SIGGSSARFGIMSDHQFNQPMCSVWASHVSHLPTNLSFIWIAPPAGTCVNMAT
	ATHRGQVIRKALAOQLCEQGAFTDVTYVHPHAEIHSDSIILRDVDSFTHQLQAPNI
	WVECNCTGECQGAIMHNAVTCFCEYPRELITTLGNTTASVLFQSGSCRFSS
	YSDPSIIVLYAKNSADWTLQEKIRAPSNVSTIHLILPEDAKGNVQFQKQNSLR
	YSGVEACWALDNILINSARQVVELESDPDVTGNMLFFPGATVVKHSCGTIESLI
	PHNGESFENFATRDVLTSTEDIOEONSEFEESOFTGMDVLGAVIGTCEGTIESLS
	MYFLKDGERRKLTPSMDTGTGYNLRFYFVWGICDPGNSHENDIILYAKIERKEHIT
	LDTLSSTKYPSLVSVINPELQTPATKFCURQNHQNNRNWAVDFHVLVPLPS
	VTQSMQIGASYMIQFSLVMGCGKITPHMDNQVLEYSTNHGLTWHLVQEELPSMP
	SCQFTSASIHASEFTQWRRVIVLLPQKTWSSATRFMSQSYTAAQDWAALDSYIG
	OOCNPMCSHGSCDHGICRCDOGYOGTECHPEAALPTIMSDFENQNGWESDQEVIG
	GEIVKPEOGCGVISSGSLYFSKAGKROLVSWDLDTSMVDYVQFYIOIGGESACNKP
	DSREGVLLQYNSNGGIOHLLAEMYFSDFSKPRVYLELPAAKTPTCTFRFWQPVF
	SEGEDQWAVDDIILISEKQKOLIPINPTLPNQYKPAFDYPMNOMSVMILANEG
	MYKNETFCATPSAMIFGKSDGDRFAVTRDLTKPGYVLQFKLNGCANQFSCTAPVL
	LOYSHDAGSMFLVKEGCPYPSAGKCGNSRELSEPTMYHTGDFEEMTRITVIPS
	LASSCTFRFRWNOESSOKNPVPGDLYVISPSPCSYSGHGDCTSGVYCFDLGTAA
	OGTCVSNVNNHMEFDFEGLKSLPWYKITGAQVGTGCTLNDGKSLYFNGPKREAR
	TVPLTRNIRLVQFYIOIGSKTSGITCIKPRTRNGLIYOVSNDNGILHLLRLDPM
	SFLEPQIISIDLDPQDAPATARFWHQPHQKHSQAALDDVLIIGNDSVSSQGTQDKF
	SCSIDLQANWYRIOGGQVDICLSMDTALIFETNIGKPRYAEWTFHVSATFFLOFM
	SMGCKSPFNSHVSQLOYSLNKKDWHLVTECVPTLVVGLMCTHYTSSITSRPNWK
	RITVYLPSTISPRFRFRWQYVAGDSWADINVLASGCPMWSGCRGICDAGRCV
	CDRGFGPYCPVVPPLPSILKDDFNGLHPDLWPBYGAERNGNLNETIKSTQSLIEK
	GEGLEMLISRLDCTNMYQVQSLFIAKSTPERSHILLOFSISGGTLWHLMDEYF
	POTNLIIFINPLPTAQTNATFRMLQYPNNGKKEEIVDDFIIDNGNNVNPVALL
	DTDFGPRDNMFYPPGGNIGLYCYSKGAPEEDSAMVVSNEGEHSITTRDLNIN
	ENTITIQFINYGCSTSSADPVRLEFSDRGATWHLLPLCYHSSSSSLCSTEHH
	PSSTVYAGTMOGWRREVYFVFGKHLICGSRFRWYQGFYPAGSQTWATIDNVIQPOC
	EMCMGQSCINGTKICIDPGYSGPTCKISTKNPDFLKDDFEGGOLDESDFKSGKP
	SRKCGILSSNNLNFNEGLRMLMTDRDLDSLHARVQFMRLGCCGKGVDPDRSQPVLL
	QYSLUNGSLWSLLQEFLEFSNNSNVGRYIALEIPLKARSGSTRLRWQSPENGHFYSPW
	VIDQILIGGNISGNTVLEDDFTTLDNRKWLHPGGTKMPVCGTGLDVFIEKASTRY

VVSTDAVNEDSFLQIDFAASCVTSDSCYAIELEYSDVLGLSWHPLVRDCLPTNVECS
RTHQLRLVTFNKNWRLITLPLPYTSQATFRFWHPAPDFKQOTQWALDINVIYGDG
CDMGSGHRCFOGNCVDEOGLGYCDPSETSLPTQLKDNFNRAFPSSQWLTIVNGGK
LSTVGAAGSALHFGSGCSRLLVTLVNLNTPAEKIOFYFMYGCLITPNNRQGVLL
EYVNGGITTWNLMEIFVDOYSKPGFVNILLPPDAKEIATFRWQPRHDJLDNDWA
IDNVLGSADORTVMDLTFSSAPVPOHSPADAGPVGRIAFDMEDKTSVNEHWL
FHDGCTVRECDSPDGMVLCGSHDREYVANTHDLTPTEGTMQKISVGCVKSEKIA
QNIHVQSTDRGVSNNLYPOCLPAEDPKCSGVSQSPSVFFPTGKWKRTIYPLPESLV
GNPVRFRYKYSMDMOWAIDNFIYLGPCLDNCRGHGDCLEQICIDPGYSGPNCYLTH
TLTKFLKERFDEEEKPDIMWSLEGSTCEGILAEIDTALYFGGTVRQAVTODLH
RGAFTLQWGRIGSNMNTSCHRPICRKEGVLLDSTGGITWTLHHEMDYOKYISVR
HDYILLPDLANTLRLMWPDEVISNGIVSVSVERAOWALDNLIGAEINPSOLVD
TDDCTSHHEWNSPVNAVRTAGCGNPSFLYWPKNKKDKKTHNALSRLEIIQPGY
MMQFKIVYGCETSGDLSHYMLEYTKDARSDSWOLVQTCPLSSNSIGCSPPFQHE
ATIIYNSVNSKKRITIQLDHVSSATQFRWIQKEETEKSQSWAIDHVIYGEACPKL
CSGHYGTGATCICDESFQDQSVFSDHSPYTKDNFESARVETANNETIQGGVIG
SCGGLAPYAGHDSLYFNGCQIQRAATPLDLTRASKIMFVLIQIGMSQTDSCSDLS
GPHAVDKVLLQVSVNNGITMHWIAHQPKDFTQAOVSYNVPLRARMKGVLLRWQOP
RHNGTGHQWALDHVEVVLVSTRKONYMMNFSROHGLRHFYNNRRSLRRYP"
BASE COUNT 3014 a 2696 c 2753 g 3116 t 1 others
ORIGIN

Query Match 75.1%; Score 263.6; DB 9; Length 11580;
Best Local Similarity 84.6%; Pred. No. 3e-69;
Matches 296; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 1 GAGCAGTGTGGCACCATCATGCATGCAATGCTGTACCTTCTGTGAGCGGTACGGCCCT 60
DB 860 GAACAGTGTGGCGGATTTATGATGCAATGCGTCACTTCTGTGAACCATATGCGCCA 919
QY 61 CGAGAGCTGACCAACATGCTCTGAACACAAACAGCATCTGCTCTCAGTTTCCATT 120
DB 920 CGAGACTGATACACAGCCCTTATACAAACAGCTTCTGCTCTCAATTTCCATT 979
QY 121 GGCTCAGGATCATGTGATTTAGTTACTCTGACCCAGCATCATCTGTGTCATCGGCAAG 180
DB 980 GGCTCAGGTTCTGCTGCTTTAGTTACTCTGACCCAGCATCATCTGTGTTATATGCGCAAG 1039
QY 181' AACAAATACCGTATGATGATTCAGCTGGAGAAATAGAGCCCTTCCAAATGTGAGCACA 240
DB 1040 AATAACTCTGCGGACTGGATTCAGCTAGAGAAATTTAGAGCCCTTCCAAATGTGAGCACA 1099
QY 241 GTCATCCACATCTGTACCTCCCGAGGAGCCAAAGGAGCGTGCAGTTCCAGTGG 300
DB 1100 ATCATCCATATCTCTACCTTCTGAGGAGCGCCAAAGGAGGAATGTCATTTTCAGTGG 1159
QY 301 AAGAGGACAGCTGCGAGTGGGTGAGGTGTATGAGGCCCTGCTGGGCCCT 350
DB 1160 AAGCAGGAAAATCTTCTGTGTAGTGAAGTGTATGAAGCCTGCTGGGCCCT 1209

RESULT 6
G30936
LOCUS SWSS2926 Eric D. Green Homo sapiens STS genomic, sequence tagged
DEFINITION site.
ACCESSION G30936
VERSION G30936.1 GI:1923209
KEYWORDS STS.
SOURCE Homo sapiens.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 11580)
Bouffard,G.G., Iyer,L.M., Idol,J.R., Braden,V.V., Cunningham,A.F.,
Weintraub,L.A., Mohr-Tidwell,R.M., Peluso,D.C., Fulton,R.S.,
Leckie,M.P. and Green,E.D.
A collection of 1814 human chromosome 7-specific STSs
Genome Res. 7 (1), 59-64 (1997)

JOURNAL
MEDLINE 97189344
PUBMED 9037602
REFERENCE 2 (bases 1 to 11580)
AUTHORS Green,E.D.

TITLE
COMMENT

Human chromosome 7 STSs (1997)
Unpublished (1997)
On Apr 3, 1997 this sequence version replaced gi:1706935.
Synonyms: RELN
GDB.DSEG: RELN
Contact: Eric D. Green
Genome Technology Branch
National Human Genome Research Institute/NIH
49 Convent Dr., MSC4431, Bldg. 49, Rm. 2A08, Bethesda, MD 20892
Tel: 3014020201
Fax: 3014024735
Email: egreen@nih.nih.gov
Primer A: CTCACCTTCTGAGGAGCGCAA
Primer B: AAGCAGTGTGCTGTGTCACCTG
STS size: 183
PCR Profile:
Presoak: 0 degrees C for 0.00 minute(s)
Denaturation: 92 degrees C for 0.17 minute(s)
Annealing: 55 degrees C for 1.00 minute(s)
Polymerization: 72 degrees C for 1.00 minute(s)
PCR Cycles: 35
Thermal Cycler: PerkinElmer 9600
Protocol:
Template: 30-100 ng
Primer: each 1 uM
dNTPs: each 200 uM
Tag Polymerase: 0.05 units/ul
Total Vol: 10 ul
Buffer: MgCl2: 1.5 mM
KCl: 100 mM
Tris-HCl: 10 mM
NH4Cl: 5 mM
pH: 8.6

This STS was developed from sequence determined by another investigator. See GenBank record: U79716 For additional information about the NHGRI chromosome 7 mapping project, see <http://www.nhgrl.nih.gov/DIR/STB/CHRT>. Also see Genomics 11:548-64 (1991) [MUID=92128937].

FEATURES
Source

gene
STS
primer_bind
primer_bind
BASE COUNT 3014 a 2696 c 2753 g 3116 t 1 others
ORIGIN

Query Match 75.1%; Score 263.6; DB 11; Length 11580;
Best Local Similarity 84.6%; Pred. No. 3e-69;
Matches 296; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 1 GAGCAGTGTGGCACCATCATGCATGCAATGCTGTACCTTCTGTGAGCGGTACGGCCCT 60
DB 860 GAACAGTGTGGCGGATTTATGATGCAATGCGTCACTTCTGTGAACCATATGCGCCA 919
QY 61 CGAGAGCTGACCAACATGCTCTGAACACAAACAGCATCTGCTCTCAGTTTCCATT 120
DB 920 CGAGACTGATACACAGCCCTTATACAAACAGCTTCTGCTCTCAATTTCCATT 979
QY 121 GGCTCAGGATCATGTGATTTAGTTACTCTGACCCAGCATCATCTGTGTCATCGGCAAG 180
DB 980 GGCTCAGGTTCTGCTGCTTTAGTTACTCTGACCCAGCATCATCTGTGTTATATGCGCAAG 1039

[illegible]

```

Qy 337 GCCTGCTGGGCCCTG 351
Db 15262 GCCTGCTGGGCCCTG 15248

RESULT 10
AC023062 183156 bp DNA linear HTG 05-MAY-2000
LOCUS
DEFINITION
Mus musculus chromosome 5 clone RP23-466D2 strain C57BL6/J, WORKING
DRAFT SEQUENCE, 8 unordered pieces.
AC023062
AC023062.2 GI:7709916
HTG; HTGS_PHASE1; HTGS_DRAFT.
Mus musculus.
Mus musculus.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 183156)
Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W.,
Bouffard, G.G., Dietrich, N.L., Eagle, W.O., Gupta, J., Ho, S.-L.,
Huang, M.C., Idol, J., Lee-Lin, S.-O., Maduro, O.L., Maduro, V.B.,
Mastrian, S.D., McCloskey, J.C., Morse, E., Ojodu, M.A., Pearson, R.,
Stantripop, S., Summers, T.J., Thomas, J.W., Thomas, P.J.,
Tiongson, E.E., Touchman, J.W., Tran, J.T., Vogt, J.L., Walker, M.A.,
Wetherby, K.D. and Green, E.D.
NISC Mouse Sequencing Initiative
Unpublished
2 (bases 1 to 183156)
Green, E.D.
Direct Submission
Submitted (08-FEB-2000) NIH Intramural Sequencing Center, 8717
GroveMont Circle, Gaithersburg, MD 20877, USA
On May 5, 2000 this sequence version replaced gi:6939141.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_mouse@nhgri.nih.gov
----- Project Information
Center project name: Y1
Center clone name: 466D02
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 175371 bases at least Q40
Consensus quality: 178474 bases at least Q30
Consensus quality: 180133 bases at least Q20
Insert size: 179000; agarose-fp
Insert size: 188000; pulse-field-gel
Insert size: 183156; sum-of-contigs
Quality coverage: 6.96x in Q20 bases; agarose-fp
Quality coverage: 6.63x in Q20 bases; pulse-field-gel
Quality coverage: 6.80x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
2374: contig of 2374 bp in length
1
2375 gap of unknown length
4490: contig of 2116 bp in length
4491 gap of unknown length
14959: contig of 10469 bp in length
14960 gap of unknown length
31072: contig of 16113 bp in length
31073 gap of unknown length
50119: contig of 19047 bp in length
gap of unknown length

```

```

* 50120 78196: contig of 28077 bp in length
* gap of unknown length
* 78197 109874: contig of 31678 bp in length
* gap of unknown length
* 109875 183156: contig of 73282 bp in length.
FEATURES
source
Location/Qualifiers
1..183156
/organism="Mus musculus"
/strain="C57BL6/J"
/db_xref="taxon:10090"
/chromosome="5"
/clone="RP23-466D2"
/clone_lib="RPC1 mouse BAC library 23"
BASE COUNT 53719 a 37080 c 37302 g 55052 t 3 others
ORIGIN
Query Match 38.5%; Score 135; DB 2; Length 183156;
Best Local Similarity 100.0%; Pred. No. 1.le-29;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 217 AGAGCCCCCTTCCAACTGTGAGCACAGTCATCCACATCTGTACCTCCCGAGGAGCCAAA 276
|||||
Db 173623 AGAGCCCCCTTCCAACTGTGAGCACAGTCATCCACATCTGTACCTCCCGAGGAGCCAAA 173682
|||||
QY 277 GGGGAGAGCGTCAGTTCCTGAGTGGAAACAGGACAGCTGCGAGTGGGTGAGTGTATGAG 336
|||||
Db 173683 GGGGAGAGCGTCAGTTCCTGAGTGGAAACAGGACAGCTGCGAGTGGGTGATGAG 173742
|||||
QY 337 GCCTGCTGGGCCCTG 351
|||||
Db 173743 GCCTGCTGGGCCCTG 173757
|||||
RESULT 11
AC095877 126130 bp DNA linear HTG 11-JUL-2002
LOCUS
DEFINITION
Rattus norvegicus clone CH230-10G13, *** SEQUENCING IN PROGRESS
***, 59 unordered pieces.
AC095877
AC095877.4 GI:21722955
HTG; HTGS_PHASE1.
Norway rat.
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 126130)
AUTHORS
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T.,
Barbaria, J., Benton, J., Blmage, K., Blankenburg, K., Bonnin, D.,
Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
DeLouth, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,
Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karlisson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louiseged, H.,
Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mel, G., Metzker, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,

```

Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwuonu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Scherer, S., Scott, G., Shen, H., Shooshtari, N., Slisson, I.,
Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 126130)
Worley, K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 126130)
Worley, K.C.
Direct Submission
Submitted (11-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 10, 2002 this sequence version replaced gi:20975941.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GDRS
Center clone name: CH230-10G13
----- Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 70934 bases at least Q40
Consensus quality: 75985 bases at least Q30
Consensus quality: 80731 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 59 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1382: contig of 1382 bp in length
* 1383 1482: gap of unknown length
* 1483 3046: contig of 1564 bp in length
* 3047 3146: gap of unknown length
* 3147 4489: contig of 1343 bp in length
* 4490 4589: gap of unknown length
* 4590 5786: contig of 1197 bp in length
* 5787 5886: gap of unknown length
* 5887 7115: contig of 1229 bp in length
* 7116 7215: gap of unknown length
* 7216 8771: contig of 1556 bp in length
* 8772 8871: gap of unknown length
* 8872 10260: contig of 1389 bp in length
* 10261 10360: gap of unknown length
* 10361 11370: contig of 1010 bp in length
* 11371 11470: gap of unknown length
* 11471 12710: contig of 1240 bp in length
* 12711 12810: gap of unknown length
* 12811 14180: contig of 1370 bp in length
* 14181 14280: gap of unknown length
* 14281 16048: contig of 1768 bp in length
* 16049 16148: gap of unknown length
* 16149 17569: contig of 1421 bp in length
* 17570 17669: gap of unknown length
* 17670 18820: contig of 1151 bp in length
* 18821 18920: gap of unknown length
* 18921 20209: contig of 1289 bp in length
* 20210 21633: contig of 1324 bp in length
* 21634 21733: gap of unknown length
* 21734 23269: contig of 1536 bp in length
* 23270 23369: gap of unknown length
* 23370 24647: contig of 1278 bp in length
* 24648 24747: gap of unknown length
* 24748 25800: contig of 1053 bp in length
* 25801 25900: gap of unknown length
* 25901 27017: contig of 1117 bp in length
* 27018 27117: gap of unknown length
* 27118 29183: contig of 2086 bp in length
* 29184 30315: contig of 1032 bp in length
* 30316 30415: gap of unknown length
* 30416 32086: contig of 1671 bp in length
* 32087 32186: gap of unknown length
* 32187 33395: contig of 1209 bp in length
* 33396 33495: gap of unknown length
* 33496 35663: contig of 2168 bp in length
* 35664 35763: gap of unknown length
* 35764 38044: contig of 2281 bp in length
* 38045 38144: gap of unknown length
* 38145 39845: contig of 1701 bp in length
* 39846 41982: contig of 2037 bp in length
* 41983 42082: gap of unknown length
* 42083 43743: contig of 1661 bp in length
* 43744 43843: gap of unknown length
* 43844 45904: contig of 2061 bp in length
* 45905 46004: gap of unknown length
* 46005 47668: contig of 1663 bp in length
* 47669 49833: contig of 2066 bp in length
* 49834 49933: gap of unknown length
* 49934 51434: contig of 1501 bp in length
* 51435 53779: contig of 2244 bp in length
* 53780 55552: contig of 1674 bp in length
* 55553 57670: contig of 2018 bp in length
* 57671 59630: contig of 1860 bp in length
* 59631 59731: gap of unknown length
* 59732 61147: contig of 1417 bp in length
* 61148 63995: contig of 2748 bp in length
* 63996 64095: gap of unknown length
* 64096 66294: contig of 2199 bp in length
* 66295 68794: gap of unknown length
* 68795 68894: gap of unknown length
* 68895 71533: contig of 2639 bp in length
* 71534 73469: contig of 1836 bp in length
* 73470 73569: gap of unknown length
* 73570 77054: contig of 3485 bp in length
* 77055 79005: contig of 1851 bp in length
* 79006 79105: gap of unknown length
* 79106 81294: contig of 2189 bp in length
* 81295 83864: contig of 2470 bp in length
83865

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```
* 83865 83964: gap of unknown length
* 83965 86039: contig of 2075 bp in length
* 86040 86139: gap of unknown length
* 86140 88646: contig of 2507 bp in length
* 88647 88746: gap of unknown length
* 88747 91823: contig of 3077 bp in length
* 91824 91923: gap of unknown length
* 91924 94369: contig of 2446 bp in length
* 94370 94459: gap of unknown length
* 94470 96886: contig of 2417 bp in length
* 96887 96986: gap of unknown length
* 96987 100679: contig of 3693 bp in length
* 100680 100779: gap of unknown length
* 100780 104171: contig of 3392 bp in length
* 104172 104271: gap of unknown length
* 104272 106760: contig of 2489 bp in length

Query Match 33.0%; Score 115.8; DB 2; Length 126130;
Best Local Similarity 91.1%; Pred. No. 7.5e-24;
Matches 123; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 217 AGAGCCCTTCCAAATGTGAGCAGTCATCCACATCTGTACCTCCCGAGGAGCCAAA 276
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 92827 AGAGCCCTTCCAAATGTGAGCAGTCATCCACATCTGTACCTCTCTGAGGAGCCAAA 92886
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 277 GGGGAGAGCGTCAGTTCCAGTGGAAACAGGACAGCCTCGAGTGGGTGAGGTGTATGAG 336
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 92887 GGGGAGAGCGTCAGTTCCAGTGGAAACAGGACAGCCTCGAGTGGGTGAGGTGTATGAG 92946
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 337 GCCTGCTGGGCCCTG 351
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 92947 GCCTGCTGGGCCCTG 92961
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
AC128022/c
LOCUS AC128022 202764 bp DNA linear HTG 19-JUL-2002
DEFINITION Rattus norvegicus clone CH230-525L20, *** SEQUENCING IN PROGRESS
***, 97 unordered pieces.
ACCESSION AC128022
VERSION AC128022.1 GI:21908605
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 202764)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,
Homsí,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulsegad,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Mosier,M., Neal,D., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu.L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojuboan,I., Rolfe,M., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tatney,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 202764)
Worley,K.C.
Direct Submission
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor-College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KAIV
Center clone name: CH230-525L20
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 113069 bases at least Q40
Consensus quality: 119970 bases at least Q30
Consensus quality: 125656 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 97 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1244: contig of 1244 bp in length
* 1245 1344: gap of unknown length
* 1345 2346: contig of 1002 bp in length
* 2347 2447: gap of unknown length
* 2447 3813: contig of 1367 bp in length
* 3814 3913: gap of unknown length
* 3914 5101: contig of 1188 bp in length
* 5102 5201: gap of unknown length
* 5202 6724: contig of 1523 bp in length
* 6725 6825: gap of unknown length
* 6825 8049: contig of 1224 bp in length
* 8049 9122: gap of unknown length
* 9123 9312: contig of 1064 bp in length
* 9313 10432: contig of 1120 bp in length
* 10433 10532: gap of unknown length
* 10533 11890: contig of 1358 bp in length
* 11891 11990: gap of unknown length
* 11991 13068: contig of 1078 bp in length
* 13069 13168: gap of unknown length
* 13169 14486: contig of 1318 bp in length
* 14487 14586: gap of unknown length
* 14587 15996: contig of 1410 bp in length
* 15997 16096: gap of unknown length

* 16097 17534: contig of 1438 bp in length
* 17335 17634: gap of unknown length
* 17635 17703: contig of 1069 bp in length
* 18704 18803: gap of unknown length
* 18804 19884: contig of 1081 bp in length
* 19885 19984: gap of unknown length
* 19985 21539: contig of 1555 bp in length
* 21540 21639: gap of unknown length
* 21640 22780: contig of 1141 bp in length
* 22781 22880: gap of unknown length
* 22881 23926: contig of 1046 bp in length
* 23927 24026: gap of unknown length
* 24027 25544: contig of 1518 bp in length
* 25545 25644: gap of unknown length
* 25645 26899: contig of 1255 bp in length
* 26900 26999: gap of unknown length
* 27000 28237: contig of 1238 bp in length
* 28238 28337: gap of unknown length
* 28338 29758: contig of 1421 bp in length
* 29759 29858: gap of unknown length
* 29859 31038: contig of 1180 bp in length
* 31039 31138: gap of unknown length
* 31139 32179: contig of 1041 bp in length
* 32180 32279: gap of unknown length
* 32280 33797: contig of 1517 bp in length
* 33797 33896: gap of unknown length
* 33897 35380: contig of 1484 bp in length
* 35381 35480: gap of unknown length
* 35481 37032: contig of 1552 bp in length
* 37033 37132: gap of unknown length
* 37133 38414: contig of 1282 bp in length
* 38415 38514: gap of unknown length
* 38515 39948: contig of 1434 bp in length
* 39949 40048: gap of unknown length
* 40049 41436: contig of 1388 bp in length
* 41437 41536: gap of unknown length
* 41537 43084: contig of 1548 bp in length
* 43085 43184: gap of unknown length
* 43185 44940: contig of 1756 bp in length
* 44941 45040: gap of unknown length
* 45041 46186: contig of 1146 bp in length
* 46187 46286: gap of unknown length
* 46287 47989: contig of 1703 bp in length
* 47990 48089: gap of unknown length
* 48090 49842: contig of 1753 bp in length
* 49843 49942: gap of unknown length
* 49943 51128: contig of 1186 bp in length
* 51129 51228: gap of unknown length
* 51229 52823: contig of 1595 bp in length
* 52824 52923: gap of unknown length
* 52924 54431: contig of 1508 bp in length
* 54332 54531: gap of unknown length
* 54532 56042: contig of 1511 bp in length
* 56043 56142: gap of unknown length
* 56143 57287: contig of 1145 bp in length
* 57288 57387: gap of unknown length
* 57388 58893: contig of 1506 bp in length
* 58894 58993: gap of unknown length
* 58994 60024: contig of 1031 bp in length
* 60025 60124: gap of unknown length
* 60125 62858: contig of 2734 bp in length
* 62859 62958: gap of unknown length
* 62959 64365: contig of 1407 bp in length
* 64366 64466: gap of unknown length
* 64466 66434: contig of 1969 bp in length
* 66435 66534: gap of unknown length
* 66535 68631: contig of 2097 bp in length
* 68632 68731: gap of unknown length
* 68732 70839: contig of 2108 bp in length
* 70840 73009: gap of unknown length
* 73010 73109: contig of 2070 bp in length
* 73110 74957: contig of 1848 bp in length

* 74958 75057: gap of unknown length
* 75058 76816: contig of 1759 bp in length
* 76817 76916: gap of unknown length
* 76917 78679: contig of 1763 bp in length
* 78680 78779: gap of unknown length
* 78780 80056: contig of 1277 bp in length
* 80057 80156: gap of unknown length
* 80157 81312: contig of 1156 bp in length
* 81313 82922: contig of 1510 bp in length
* 82923 83023: gap of unknown length
* 83023 84180: contig of 1158 bp in length
* 84181 84280: gap of unknown length
* 84281 85849: contig of 1569 bp in length
* 85850 85949: gap of unknown length
* 85950 87612: contig of 1663 bp in length
* 87613 87712: gap of unknown length

Query Match 33.0%; Score 115.8; DB 2; Length 202764;
Best Local Similarity 91.1%; Pred. No. 8.3e-24;
Matches 123; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 217 AGAGCCCTTCCAATGTGAGCAGATCCACATCCTGTACCTCCCGAGAGGACAAA 276
|||||
Db 187315 AGAGCCCTTCCAACGTAGACACCATCATCCATCCTCTTCTGAGGAGGACAAA 187256
|||||

QY 277 GGGGAGAGCGTGCAGTCCAGTGGAAACAGACAGCGTGGAGTGGGTGATGAG 336
|||||
Db 187255 GGGGAGAACGTGCAGTCCAGTGGAAACAGACAGCGTGGGTGATGAGTACGAA 187196
|||||

QY 337 GCCTGCTGGGCCCTG 351
|||||
Db 187195 GCCTGCTGGGCCCTG 187181
|||||

RESULT 13
AC124933/c
LOCUS Rattus norvegicus clone CH230-466N17, *** SEQUENCING IN PROGRESS
DEFINITION ***
ACCESSION AC124933
VERSION AC124933.2 GI:21952716
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 183641)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Alsbrooks,S.L., Anaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carton,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,K., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J.J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homsif,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jollivet,S., Joudah,S.,
Karlssoon,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,

```

Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokkwo,S., Ogih,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojebokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherrer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I.,
Sodergren,E., Sonaika,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansley,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wlarczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 183641)
Worley,K.C.
Direct Submission
Submitted (20-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 183641)
Worley,K.C.
Direct Submission
Submitted (31-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 24, 2002 this sequence version replaced gi:21490053.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KAVI
Center clone name: CH230-466N17
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 111724 bases at least Q40
Consensus quality: 120970 bases at least Q30
Consensus quality: 128458 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 88 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1107: contig of 1107 bp in length
1108 1207: gap of unknown length
1208 1260: contig of 1253 bp in length
2461 2560: gap of unknown length
2561 4004: contig of 1444 bp in length
4105 4104: gap of unknown length
4105 5382: contig of 1278 bp in length
5383 5482: gap of unknown length
5483 6874: contig of 1392 bp in length
6875 6974: gap of unknown length
6975 8071: contig of 1097 bp in length
8072 8171: gap of unknown length
8172 9478: contig of 1307 bp in length
9479 9578: gap of unknown length
9579 11061: contig of 1483 bp in length
11062 11161: gap of unknown length

11162 12742: contig of 1581 bp in length
12743 12842: gap of unknown length
12843 14504: contig of 1662 bp in length
14505 14604: gap of unknown length
14605 15734: contig of 1130 bp in length
15735 15834: gap of unknown length
15835 16958: contig of 1124 bp in length
16959 17059: gap of unknown length
17059 18705: contig of 1646 bp in length
18705 18805: gap of unknown length
18805 20420: contig of 1616 bp in length
20421 20521: gap of unknown length
20521 21649: contig of 1128 bp in length
21649 23177: contig of 1429 bp in length
23178 23277: gap of unknown length
23278 24761: contig of 1483 bp in length
24761 24861: gap of unknown length
24861 26759: contig of 1899 bp in length
26760 26859: gap of unknown length
26860 28399: contig of 1440 bp in length
28399 28400: gap of unknown length
28400 29592: contig of 1193 bp in length
29593 30870: gap of unknown length
30871 30970: contig of 1178 bp in length
30971 32642: contig of 1672 bp in length
32643 32743: gap of unknown length
32743 34386: contig of 1644 bp in length
34387 34486: gap of unknown length
34487 36064: contig of 1578 bp in length
36065 37398: contig of 1234 bp in length
37399 37499: gap of unknown length
37499 39608: contig of 2110 bp in length
39609 41480: gap of unknown length
41481 41580: contig of 1772 bp in length
41581 42759: contig of 1179 bp in length
42760 42859: gap of unknown length
42860 44711: contig of 1852 bp in length
44712 44811: gap of unknown length
44812 46208: contig of 1397 bp in length
46209 46309: gap of unknown length
46309 47767: contig of 1459 bp in length
47768 47868: gap of unknown length
47868 49227: contig of 1360 bp in length
49228 49327: gap of unknown length
49328 50797: contig of 1470 bp in length
50798 51928: contig of 1031 bp in length
51929 52028: gap of unknown length
52029 53105: contig of 1077 bp in length
53106 53205: gap of unknown length
53206 54546: contig of 1341 bp in length
54547 55820: contig of 1174 bp in length
55821 55920: gap of unknown length
55921 57635: contig of 1715 bp in length
57636 58942: contig of 1207 bp in length
58943 59042: gap of unknown length
59043 60341: contig of 1299 bp in length
60342 60441: gap of unknown length
60442 61693: contig of 1252 bp in length
61694 61793: gap of unknown length
61794 63315: contig of 1522 bp in length
63316 64571: contig of 1156 bp in length
64572 64672: gap of unknown length
64672 66083: contig of 1412 bp in length
66084 66183: gap of unknown length
66184 67516: contig of 1333 bp in length

```

```

* 67517 67616: gap of unknown length
* 67617 69820: contig of 2204 bp in length
* 69821 69920: gap of unknown length
* 69921 71261: contig of 1341 bp in length
* 71262 72775: gap of unknown length
* 72776 72875: contig of 1414 bp in length
* 72876 75298: gap of unknown length
* 75299 75398: contig of 2423 bp in length
* 75399 77094: gap of unknown length
* 77095 77194: contig of 1696 bp in length
* 77195 78834: gap of unknown length
* 78835 78934: contig of 1640 bp in length
* 78935 79988: gap of unknown length
* 79989 80089: contig of 1055 bp in length
* 80090 81774: gap of unknown length
* 81775 81874: contig of 1685 bp in length
* 81875 83255: gap of unknown length
* 83256 83255: contig of 1381 bp in length

Query Match 32.58; Score 114.2; DB 2; Length 183641;
Best Local Similarity 90.4%; Pred. No. 2.5e-23;
Matches 122; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 217 AGAGCCCTTCCATGTGAGCAGACATCCATCATCTGACCTCCCGAGGAGCCAAA 276
|||||
Db 143722 AGAGCCCTTCCACGTAAGCACCATCATCTGACCTCCCGAGGAGCCAAA 143663
|||||

Qy 277 GGGGAGCGTGCAGTTCCTGGAACAGAGCAGCGCTGCGAGTGGGTGAGTGTATGAG 336
|||||
Db 143662 GGGGAGAGCGTGCAGTTCCTGGAACAGAGCAGCGCTGCGTGGGTGAGTGTATGAG 143603
|||||

Qy 337 GCCTGCTGGGCCCTG 351
|||||
Db 143602 GCCTGCTGGGCCCTG 143588

RESULT 14
HSAC000121/c
LOCUS HSAC000121 93163 bp DNA linear PRI 31-JAN-1997
DEFINITION Human BAC clone RG249A12 from 7q22, complete sequence.
ACCESSION AC000121
VERSION AC000121.1 GI:1809222
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Lamar,B., Wamsley,P. and Gibson,A.
1 (bases 1 to 93163)
The sequence of H. sapiens BAC clone RG249A12
Unpublished (1997)
REFERENCE
1 Waterston,R.
2 (bases 1 to 93163)
Direct Submission
Submitted (31-JAN-1997)
JOURNAL
Genome Sequencing Center
Department of Genetics, Washington University
St. Louis, MO 63108, USA
http://genome.wustl.edu/gsc
e-mail: sapiens@watson.wustl.edu

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

```

SOURCE INFORMATION:
This clone is from a release of the human BAC library. The library
contains cloned DNA from a human sperm. For references see:
Shizuya et al., Proc. Natl. Acad. Sci. 89:8794-8797 (1992); Kim et
al., Genomics 34:213-218 (1996).
VECTOR: pBELO
Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the right is H_RG352J05. Actual start of
this clone is at base position 1 of H_RG249A12; actual end is at
93163 of H_RG249A12. This clone is part of an unanchored island,
orientation is unknown.

This clone contains STS SWSS2926 (NID:gl706935).

FEATURES
Source
Location/Qualifiers
1. 93163
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q22"
/clone="RG249A12"
/complement(1011..1309)
/rpt_family="ALU"
repeat_region 2887..3178
repeat_region 2887..3178
misc_feature 5965..6163
/notes="similar to C2H2-type zinc fingers PROSITE; PS00028"
repeat_region complement(7118..7170)
repeat_region 9584..12339
repeat_region 10207..11298
exon
/notes="GRAIL prediction, score = 85"
/feature="not_experimental"
complement(10583..11001)
repeat_region 12423..12714
repeat_region 12715..12734
repeat_region 13365..13412
repeat_region 14286..14309
repeat_region 15942..15963
repeat_region 15942..15963
gene complement(<17055..>84441)
/notes="WUGSC:H_RG249A12.1"
CDS complement(join(<17055..17200,32792..33032,35849..35945,
58079..58130,63050..63146,84365..>84441))
/notes="WUGSC:H_RG249A12.1"
/notes="extracellular matrix protein; 97% similarity to
mouse reelin U24703 (PID:902487), partial CDS"
/codon_start=1
/evidence-not_experimental
/protein_id="AAB46357.1"
/db_xref="GI:1809223"
/translation="EIHSDSIILRDFDSYHQLQNPNIWECNCCETGBOCGAIMHG
NATKCEPYGPRELITGLNTTASVLPQSIGSGCRFSYSDPSIIVLYAKNSADMI
QLEKIRAPNSVTIIHILYPEDAKGNQVQKQENLRVGEVYECACWALDNLIIINS
AHRQVYLEDSDPDVTGNLFFPGATVKHSCQSDGNSIYPHNEGSEFNPFATTRVDL
STEDIQEOWSEEFESQPTG"
complement(17055..17252)
exon
/notes="WUGSC:H_RG249A12.1"
/notes="GRAIL prediction, score = 100"
/evidence-not_experimental
repeat_region 21484..21521
/rpt_family="L1"
repeat_region complement(21962..21990)
repeat_region 24430..24466
/rpt_family="L1"
repeat_region 24430..24466

```

```
repeat_region /rpt_family="L1"
24613..24654
repeat_region /rpt_family="L1"
25141..25180
repeat_region /rpt_family="L1"
26470..26761
misc_feature /rpt_family="ALU"
/genes="WUGSC:H_RG249A12.1"
complement(26762..26936)
repeat_region /note="match to Human cDNA, 5' end H56746 (NID:1005390)"
26937..26989
exon /rpt_family="L1"
27107..27238
/notes="GRAIL prediction, score = 82"
/evidence=not_experimental
repeat_region 27993..28465
repeat_region /rpt_family="L1"
28476..28496
repeat_region /rpt_family="L1"
complement(29663..29903)
exon /rpt_family="ALU"
/genes="WUGSC:H_RG249A12.1"
/notes="GRAIL prediction, score = 90"
/evidence=not_experimental
complement(33462..33752)
repeat_region /rpt_family="ALU"
37975..37994
repeat_region /rpt_family="L1"
40278..40311
repeat_region /rpt_family="L1"
complement(41921..42075)
repeat_region /rpt_family="ALU"
complement(42447..42734)
exon /rpt_family="ALU"
43126..43185
/notes="GRAIL prediction, score = 95"
/evidence=not_experimental
complement(43411..43680)
repeat_region /rpt_family="ALU"
45704..45991
misc_feature /rpt_family="ALU"
46064..46434
/notes="match to Human cDNA H50097, 5' end (NID:g1012929)"
repeat_region complement(46604..46897)
misc_feature /rpt_family="ALU"
complement(47023..47412)
/genes="WUGSC:H_RG249A12.1"
/notes="match to Human cDNA, 3' end H59444 (NID:g1012276)"
58125..58163
repeat_region /rpt_family="L1"
58604..58758
repeat_region /rpt_family="L1"
59743..59765
repeat_region /rpt_family="L1"
60210..60516
exon /rpt_family="ALU"
complement(63051..63146)
/genes="WUGSC:H_RG249A12.1"
/notes="GRAIL prediction, score = 100"
/evidence=not_experimental
64242..64284
repeat_region /rpt_family="L1"
64445..64496
repeat_region /rpt_family="L1"
complement(67753..68011)
repeat_region /rpt_family="L1"
67986..68889
exon /rpt_family="L1"
complement(70725..70841)
/genes="WUGSC:H_RG249A12.1"
/notes="GRAIL prediction, score = 82"
/evidence=not_experimental
```

```
repeat_region 71911..71947
/rpt_family="L1"
complement(73631..73919)
repeat_region /rpt_family="ALU"
75073..75092
repeat_region /rpt_family="L1"
76743..82366
exon /rpt_family="L1"
78326..79402
/notes="GRAIL prediction, score = 83"
/evidence=not_experimental
80586..81548
/notes="GRAIL prediction, score = 82"
/evidence=not_experimental
complement(80962..81380)
repeat_region /rpt_family="L1"
complement(82581..82608)
/rpt_family="L1"
83877..84447
misc_feature

Query Match 28.1%; Score 98.8; DB 9; Length 93163;
Best Local Similarity 83.6%; Pred. No. 1.1e-18;
Matches 112; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 217 AGAGCCCTTCCAAATGTGACACAGTCATCCACATCCTGTACCTCCCGAGGACCAAA 276
|||||
Db 33034 AGAGCCCTTCCAAATGTGACACAAATCATCCATATCCTCTACCTCTCTGAGGACGCCAA 32975
|||||

QY 277 GGGGAGAGCGTGCAGTTCAGTGGAAACAGGACAGCGCTGCGAGTGGTGGTGTATGAG 336
|||||
Db 32974 GGGGAGAATGTCCAAATTTTCAGTGGAAAGCAGGAAAATCTTCGTAGTGAAGTGTATGAA 32915
|||||

QY 337 GCCTGCTGGGCCCT 350
|||||
Db 32914 GCCTGCTGGGCCCT 32901
|||||
```

RESULT 15
AC041023/c

AC041023 Homo sapiens clone RP11-356G3, WORKING DRAFT SEQUENCE, 20 unordered pieces.
LOCUS AC041023.2 GI:8016758
DEFINITION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 185996)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens, clone RP11-356G3
JOURNAL Unpublished
AUTHORS

REFERENCE 2 (bases 1 to 185996)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F., Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczkv,J., Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrum,J., Meneus,L., Mihoval,T., Miranda,C., Mlenga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Olivier,T.M., Oliver,J., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,

THIS PAGE BLANK (USPTO)

Result No.	Score	Query		Length	DB	ID	Description
		Match					
1	351	100.0	351	10	US-09-897-438B-1		Sequence 1, Appli
2	263.6	75.1	11580	10	US-09-880-107-3436		Sequence 3436, Ap
C 3	98.8	28.1	499	10	US-09-864-761-14497		Sequence 14497, A
C 4	81.4	23.2	214	10	US-09-864-761-31037		Sequence 31037, A
C 5	53.8	15.3	413	10	US-09-864-761-11808		Sequence 11808, A
C 6	46.2	13.2	187	10	US-09-864-761-28379		Sequence 28379, A
C 7	31.6	9.0	168	10	US-09-878-574-14485		Sequence 14485, A
8	31.4	8.9	670	10	US-09-828-644-8		Sequence 8, Appli
9	31.4	8.9	1536	12	US-10-050-726-1		Sequence 1, Appli
10	31.4	8.9	1860	10	US-09-828-644-9		Sequence 9, Appli
C 11	31.2	8.9	417	10	US-09-922-217-605		Sequence 605, App
C 12	31.2	8.9	417	10	US-09-833-263-605		Sequence 605, App
C 13	31.2	8.9	2532	12	US-10-124-598-4		Sequence 4, Appli
14	30.6	8.7	2298	10	US-09-476-242-22		Sequence 22, Appli
15	30.6	8.7	2298	10	US-09-476-242-23		Sequence 23, Appli
16	30.6	8.7	2298	10	US-09-476-242-24		Sequence 24, Appli
17	30.6	8.7	2310	10	US-09-476-242-3		Sequence 3, Appli
18	30.6	8.7	2310	10	US-09-476-242-21		Sequence 21, Appli
19	30.6	8.7	2316	10	US-09-476-242-4		Sequence 4, Appli

Db	241	GTGATTCACATCTGTACCTCCCGAGGAAGCCAAAGGGAGAGCGTGCAGTTCCAGTGG	300
Qy	301	AAACAGGACACCGCTGCAGTGGGTGAGTGCTATGAGCGCCTGCTGGGCCCTG	351
Db	301	AAACAGGACACCGCTGCAGTGGGTGAGTGCTATGAGCGCCTGCTGGGCCCTG	351

RESULT 2
 US-09-880-107-3436
 ; Sequence 3436, Application US/09880107
 ; Patent No. US20020142981A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Horne, Darci T.
 ; APPLICANT: Vockley, Joseph G.
 ; APPLICANT: Scherf, Uwe
 ; APPLICANT: Gene Logic, Inc.
 ; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
 ; FILE REFERENCE: 44921-5028-WO
 ; CURRENT APPLICATION NUMBER: US/09/880,107
 ; PRIOR FILING DATE: 2001-06-14
 ; PRIOR APPLICATION NUMBER: US 60/211,379
 ; PRIOR FILING DATE: 2000-06-14
 ; PRIOR APPLICATION NUMBER: US 60/237,054
 ; PRIOR FILING DATE: 2000-10-02
 ; NUMBER OF SEQ ID NOS: 3950
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3436
 ; LENGTH: 11580
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U97916
 US-09-880-107-3436

Query Match	75.1%	Score 263.6	DB 10	Length 11580
Best Local Similarity	84.6%	Pred. No. 9.4e-77		
Matches 296	Conservative 0	Mismatches 54	Indels 0	Gaps 0
QY	1	GAGCAGTGTGGCACCATCATGCATGGCAATCGCTGCACCTTCCTGTGAGCCCGTACGGCCCT	60	
Db	860	GAACAGTGTGGCGGATATCATGTGGCAATGCCGTGACCTTCCTGTGTAACCATATGGGCCA	919	
QY	61	CGAGAGCTGACCCACATCGCTGAACACAAACACAGCATCTGCCTCCAGTTCCTCCATT	120	
Db	920	CGAAGCTGATTACCCACAGGCTTAATACACACACAGCTTCCTGCTCCCAATTTCCATT	979	
QY	121	GGGTACAGATCATGTCAATTAGTTACTCTGACCCACAGCATCACTGTGTCATACGCCAAG	180	
Db	980	GGGTACAGGTTCATGTCCTTTAGTTATTCAGACCCACAGCATCATCGTTTATATGCCAAG	1039	
QY	181	AACAATACCGCTGATTCGGATTCAGCTCGAGAAAAATTAGAGCCCTTCCAATGTCAGACACA	240	
Db	1040	ATAACTCTCGGACTGGATTCAGTACAGAGAAATTTAGAGCCCTTCCAATGTCAGACACA	1099	
QY	241	GTATCCACATCTCTGTACCTCCCGAGGAAAGCCAAAGGGAGAGCGTGTGCAATTCACGTGG	300	
Db	1100	ATCATCCATATCTCTTACCTTCCTGAGGACGCCAAAGGGAGAAATGTCCAATTTCACTGG	1159	
QY	301	AAACAGGACAGCCTGCAGCTGGGTAGGTGTATGAGGCCCTGCTGGGCCCT	350	
Db	1160	AACGAGGAAATCTTCGTGTAGGTGAAGTGTATGAAGCCTGCTGGGCCT	1209	

RESULT 3
US-09-864-761-14497/c
: Sequence 14497, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng

```

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 14497
; LENGTH: 499
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC000121.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.62
; JS-09-864-761-14497

```

	Query Match	28.1%	Score 98.8;	DB 10;	Length 499;
	Best Local Similarity	83.6%	Pred. No. 4.4e-23;		
	Matches 112;	Conservative	0;	Mismatches 22;	Indels 0;
					Gaps 0;
QY	217	AGAGCCCTTCCAAATGTGAGCACAGTATCCACATCTGTACCTCCCGAGAGGCCAAA	276		
Db	487	AGAGCCCTTCCAAATGTGAGCACAAATCATCATCTCTACCTCTCTGAGGACGCCAAA	428		
QY	277	GGGAGAGCGTGCAGTTCCTAGTGGAAACAGACACAGCCTGCCAGTGGGTGAGGTGTATCAG	336		
Db	427	GGGAGAGAAATTCCAATTCAGTGGAGCAGGAGAAATCTTCGTGTAGGTGAAGTGTATGAA	368		
QY	337	GCCTGCTGGGCCCT	350		
Db	367	GCCTGCTGGGCCCT	354		

RESULT 4
US-09-864-761-31037/c
; Sequence 31037, Application US/09864761
; Patent No. US20020048763A1

```

: GENERAL INFORMATION:
: APPLICANT: Penn, Sharron G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
: FILE REFERENCE: Aeomica-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 31037
: LENGTH: 214
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC000121.1
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.62
: OTHER INFORMATION: EST_HUMAN HIT: BE38667.1, EVALUE 1.00e-76
: OTHER INFORMATION: NT HIT: g14826977, EVALUE 1.00e-117
: OTHER INFORMATION: SWISSPROT HIT: O14063, EVALUE 1.20e+00
US-09-864-761-31037

Query Match      23.2%; Score 81.4; DB 10; Length 214;
Best Local Similarity 81.7%; Pred. No. 1.6e-17;
Matches 94; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 236 GCACAGTCATCCATCCTGTACCTCCCGAGGAGCCAAAGGGGAGCGTGCAGTTC 295
Db 214 GCACATCATCCATCCTCTACCTTCTGTAGGACGCCAAAGGGGAGGAATGTCAAATTTC 155

Qy 296 AGTGAACACAGCAGACGCTGCGAGTGGGTGAGGTGTATGAGGCCTGTCTGGGCCCT 350
Db 154 AGTGAACAGCAAAATCTTCGTGTAGGTGAAGTGTATGAGCCCTGTCTGGGCCCT 100

```

```

RESULT 5
US-09-864-761-11808/c
: Sequence 11808, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharron G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
: TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
: FILE REFERENCE: Aeomica-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 11808
: LENGTH: 413
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC000121.1
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.55
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.57
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.51
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.44
US-09-864-761-11808

Query Match      15.3%; Score 53.8; DB 10; Length 413;
Best Local Similarity 79.0%; Pred. No. 2.8e-08;
Matches 64; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 1 GAGCAGTGTGGCACCACATCATGCGCAATGCTGTACCTTCTGTGAGCGCTACGGCCT 60
Db 323 GAACAGTGTGGCGCATATGCGCATGCGCAATGCGCTTCTGTGAAACCATATGGCCCA 264

Qy 61 CGAGAGCTGACCACCATGC 81

```



```
Query Match      8.9%; Score 31.4; DB 10; Length 670;
Best Local Similarity 47.3%; Pred. No. 0.85;
Matches 95; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 10 GGCACCATCATGATGCAATGCTGTGCACCTTCTGTGAGCGGTACGGCCCTCGAGAGCTG 69
    |||||
Db 5 GTCTCATCATGATGCTGTGCTGTCATCTGTGCTGGGGCCCTACTGCTTCCCTGGTGTG 64
    |||||

QY 70 ACCACCATGTCCTGAACAACAACAAGCATCTGTCTCCAGTTCCTCATTTGGGTACGGA 129
    |||||
Db 65 CTGGCCGCCGCCGAGGAGCCAGACCATGAGCCGCCCTCGCTCCTCAGCGTGGTGCC 124
    |||||

QY 130 TCATGTCGATTTAGTTACTCTGACCCAGCATCATCTGTGTACGACCAAGACAATACC 189
    |||||
Db 125 GTCTGGCTGACCTGGGCAATGGGCGCATCAACCTGTCTATCAGCCATCCGCAATCCC 184
    |||||

QY 190 GCTGATTGGATTGAGCTGGAG 210
    |||||
Db 185 AACATTTCGATGCTCTTAGG 205
    |||||

RESULT 9
US-10-050-726-1
; Sequence 1, Application us/10050726
; Patent No. US20020151697A1
; GENERAL INFORMATION:
; APPLICANT: Patirana, Marie Sudam
; TITLE OF INVENTION: DNA ENCODING ORPHAN SNORF68 RECEPTOR
; FILE REFERENCE: 60795
; CURRENT APPLICATION NUMBER: US/10/050,726
; CURRENT FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 09/466,570
; PRIOR FILING DATE: 1999-12-17
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 1
; LENGTH: 1536
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-050-726-1

Query Match      8.9%; Score 31.4; DB 12; Length 1536;
Best Local Similarity 47.3%; Pred. No. 1.3;
Matches 95; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 10 GGCACCATCATGATGCAATGCTGTGCACCTTCTGTGAGCGGTACGGCCCTCGAGAGCTG 69
    |||||
Db 1022 GTCTCATCATGATGCTGTGCTGTCATCTGTGCTGGGGCCCTACTGCTTCCCTGGTGTG 1081
    |||||

QY 70 ACCACCATGTCCTGAACAACAACAAGCATCTGTCTCCAGTTCCTCATTTGGGTACGGA 129
    |||||
Db 1082 CTGGCCGCCGCCGAGGAGCCAGACCATGAGCCGCCCTCGCTCCTCAGCGTGGTGCC 1141
    |||||

QY 130 TCATGTCGATTTAGTTACTCTGACCCAGCATCATCTGTCTACGACCAAGACAATACC 189
    |||||
Db 1142 GTCTGGCTGACCTGGGCAATGGGCGCATCAACCTGTCTATCAGCCATCCGCAATCCC 1201
    |||||

QY 190 GCTGATTGGATTGAGCTGGAG 210
    |||||
Db 1202 AACATTTCGATGCTCTTAGG 1222
    |||||

RESULT 10
US-09-828-644-9
; Sequence 9, Application us/09828644
; Patent No. US20020015998A1
; GENERAL INFORMATION:
; APPLICANT: Vogeli, Gabriel
; TITLE OF INVENTION: No. US20020015998A1e1 G Protein-Coupled Receptors
; FILE REFERENCE: 00196US1
; CURRENT APPLICATION NUMBER: US/09/828,644
; CURRENT FILING DATE: 2001-04-06
```

```
; PRIOR APPLICATION NUMBER: 60/195,150
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,099
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,151
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,148
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,093
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,098
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/230,149
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 1860
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-828-644-9

Query Match      8.9%; Score 31.4; DB 10; Length 1860;
Best Local Similarity 47.3%; Pred. No. 1.4;
Matches 95; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 10 GGCACCATCATGATGCAATGCTGTGCACCTTCTGTGAGCGGTACGGCCCTCGAGAGCTG 69
    |||||
Db 1200 GTCTCATCATGATGCTGTGCTGTCATCTGTGCTGGGGCCCTACTGCTTCCCTGGTGTG 1259
    |||||

QY 70 ACCACCATGTCCTGAACAACAACAAGCATCTGTCTCCAGTTCCTCATTTGGGTACGGA 129
    |||||
Db 1260 CTGGCCGCCGCCGAGGAGCCAGACCATGAGCCGCCCTCGCTCCTCAGCGTGGTGCC 1319
    |||||

QY 130 TCATGTCGATTTAGTTACTCTGACCCAGCATCATCTGTCTACGACCAAGACAATACC 189
    |||||
Db 1320 GTCTGGCTGACCTGGGCAATGGGCGCATCAACCTGTCTATCAGCCATCCGCAATCCC 1379
    |||||

QY 190 GCTGATTGGATTGAGCTGGAG 210
    |||||
Db 1380 AACATTTCGATGCTCTTAGG 1400
    |||||

RESULT 11
US-09-922-217-605/c
; Sequence 605, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Xuqiu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922,217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 605
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-922-217-605
```

```

Query Match      8.9%; Score 31.2; DB 10; Length 417;
Best Local Similarity 50.7%; Pred. No. 0.79;
Matches 75; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 195 TTGATTTCAGCTGAGAGAAATTAGAGCCCTTCCAAATGTGAGCAGACATCCACATCCT 254
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 339 TTGATTTCAGAGAGCCCAAAATCTAGTGTGGCCAAAGTGCCCACTCTGAGCAAACTCTT 280
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 255 GTACCTCCCGAGGAGCAAGGAGAGCGTGCGAGTTCAGTGGAAACAGGACAGCCT 314
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 279 GGAATGATGGAGAGTCCAGAGTTTACAAGAGAAGATCGACGAGAGACACACAGCCA 220
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 315 GCGAGTGGGTGAGTGTATGAGGCGCTGC 342
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 219 GCAGGCCCGCAGGCTGAGGCGCCCTGC 192
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
US-09-833-263-605/c
; Sequence 605, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stoilk, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 605
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-833-263-605

Query Match      8.9%; Score 31.2; DB 10; Length 417;
Best Local Similarity 50.7%; Pred. No. 0.79;
Matches 75; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 195 TTGATTTCAGCTGAGAGAAATTAGAGCCCTTCCAAATGTGAGCAGACATCCACATCCT 254
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 339 TTGATTTCAGAGAGCCCAAAATCTAGTGTGGCCAAAGTGCCCACTCTGAGCAAACTCTT 280
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 255 GTACCTCCCGAGGAGCAAGGAGAGCGTGCGAGTTCAGTGGAAACAGGACAGCCT 314
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 279 GGAATGATGGAGAGTCCAGAGTTTACAAGAGAAGATCGACGAGAGACACACAGCCA 220
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 315 GCGAGTGGGTGAGTGTATGAGGCGCTGC 342
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 219 GCAGGCCCGCAGGCTGAGGCGCCCTGC 192
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
US-10-124-598-4
; Sequence 4, Application US/10124598
; Patent No. US20020119526A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Lindemeyer, Juergen
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor
; FILE OF INVENTION: Involved in Sensory Transduction
; FILE REFERENCE: 02307E-088720US
; CURRENT APPLICATION NUMBER: US/10/124,598
; CURRENT FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/361,631
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/112,747

```

```

; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2532
; TYPE: DNA
; ORGANISM: Mus sp.
; FEATURE:
; OTHER INFORMATION: mouse G-protein coupled receptor (GPCR) B4
; OTHER INFORMATION: nucleotide sequence
US-10-124-598-4

Query Match      8.9%; Score 31.2; DB 12; Length 2532;
Best Local Similarity 47.4%; Pred. No. 1.9;
Matches 93; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 146 ACTCTGACCCAGCATCTGTGTATACGCCAAGACAAATACCGTGTATGATTCAGC 205
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 464 ACATTCTCTCTACTTCTCGTCCACAGTGCATATAGCGCCATCACCGACAAGCTGC 523
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 206 TGGAGAAATTTAGAGCCCTTCCAAATGTGAGCAGCATCATCCACATCTCTACTCCCG 265
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 524 AAGACAAGCGCGCTTCCCTGCCATGCTGGCACTGTGCCAGCGGCCACCCACATCG 583
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 266 AGGAAGCCAAAGGGAGAGCGTGCAGTTCAGTGGAAACAGAGACAGCTCCGAGTGGTG 325
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 584 AGGCATGGTGAACATGATGGTTCACTTCCAGTGAAGTGGTGGTCTGCTGTGAGCG 643
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 326 AGGTGTATGAGGCGCTG 341
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 644 ATGACGATTATGGCCG 659
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
US-09-476-242-22
; Sequence 22, Application US/09476242
; Patent No. US20020146683A1
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: MARTIN, Eric
; APPLICANT: HARTOG, Karin
; TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
; FILE REFERENCE: 1605.002
; CURRENT APPLICATION NUMBER: US/09/476,242
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 2298
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Vall120-Ile201;
; OTHER INFORMATION: Ile424-Ala433
US-09-476-242-22

Query Match      8.7%; Score 30.6; DB 10; Length 2298;
Best Local Similarity 52.8%; Pred. No. 2.8;
Matches 66; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 188 CCGCTGATTGGATTGAGTTCAGCTGGAGAAAATTAGAGCCCTTCCAAATGTGAGCAGTATCC 247
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 518 CCGTGTGTAGACCCAGCTGCTGCTGAACGGCAGCCTGGCCGAGGAGGGCGTGGTATCC 577
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 248 ACATCTGTACCTCCCGGAGGAGCCAAAGGGGAGAGCGTGCAGTTCAGTGGAAACAGG 307
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 578 GCAGCGAGAACTTCAACCGAAGCCCAAGCATCATCTGTCAGCTGAAGGAGCGGTGG 637
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 308 ACAGC 312
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 638 AGATC 642
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


RESULT 15
US-09-476-242-23
; Sequence 23 Application US/09476242
; Patent No. US20020146683A1
; GENERAL INFORMATION:
; APPLICANT: BARNETT, Susan
; APPLICANT: HARTOG, Karin
; APPLICANT: MARTIN, Eric
; TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES
; FILE REFERENCE: 1605.002
; CURRENT APPLICATION NUMBER: US/09/476,242
; CURRENT FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 2298
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: Val120-Ile201B; Ile424-Ala433
US-09-476-242-23

	Query Match	8.7%	Score 30.6;	DB 10;	Length 2298;
	Best Local Similarity	52.8%	Pred. No. 2.8;		
	Matches	66;	Conservative 0;	Mismatches 59;	Indels 0; Gaps 0;
Qy	188	CCGCTGATTGGATTTCAGCTGGAGAAATTTAGAGCCCTTCCAAATGTGAGCACAGTCATCC	247		
Db	518	CCGTGGTGAGCACCCAGCTGCTGTAACGGCAGCCTGGCGGAGGGCGTGTGATCC	577		
Qy	248	ACATCTGTACTCCCGAGGAAGCCAAAGGGGAGAGCGTCCAGTTCAGTGGAAACAGG	307		
Db	578	GCACGGAGAACTTCACCGACAACGCCCAAGACCACATCATGTGTCAGCTGAAGGAGAGCGTGG	637		
Qy	308	ACAGC	312		
Db	638	AGATC	642		

Search completed: November 6, 2002, 16:43:11
Job time : 91 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 6, 2002, 13:57:21 ; Search time 68 Seconds
(without alignments)
1582.993 Million cell updates/sec

Title: US-09-897-438B-1
Perfect score: 351
Sequence: 1 gacagtggtggaccatcat.....atgagggcctgctggggccctg 351

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	351	100.0	11673	4	US-09-334-220-3
2	263.6	75.1	11580	4	US-09-334-220-4
3	35.6	10.1	3344	1	US-07-718-575-13
4	35.6	10.1	3344	1	US-08-481-206-13
5	35.6	10.1	3344	2	US-08-486-269A-13
6	35.6	10.1	3619	4	US-08-377-503-1
7	35.6	10.1	3619	4	US-08-178-019-1
8	31.2	8.9	2532	4	US-09-361-631-4
9	29.8	8.5	2329	4	US-09-411-977-1
10	29.6	8.4	1176	4	US-09-372-422A-25
11	29.6	8.4	5020	3	US-08-938-291A-3
12	29.4	8.4	2657	4	US-09-488-671-3
13	29	8.3	1011	4	US-09-641-806-3
14	29	8.3	1026	4	US-09-641-806-1
15	29	8.3	2095	1	US-08-336-408B-3
16	29	8.3	2095	5	PCT-US91-00399-3
17	28.8	8.2	2565	4	US-08-985-492-7
18	28.6	8.1	935	4	US-09-221-017B-331
19	28.6	8.1	1068	5	PCT-US91-00899-13
20	28.6	8.1	1215	5	PCT-US91-00899-5
21	28.6	8.1	1488	1	US-07-914-281-9
22	28.6	8.1	1488	1	US-08-393-246-9
23	28.6	8.1	1488	1	US-08-525-058A-9
24	28.6	8.1	1488	2	US-08-696-731-9
25	28.6	8.1	1488	4	US-09-042-531-9
26	28.6	8.1	2134	2	US-08-483-151-3
27	28.6	8.1	2134	5	PCT-US96-06427-3

C 28	28.6	8.1	2175	4	US-08-482-073-9	Sequence 9, Appli
C 29	28.6	8.1	2861	4	US-08-482-073-10	Sequence 10, Appli
C 30	28.6	8.1	3647	1	US-07-914-281-7	Sequence 7, Appli
C 31	28.6	8.1	3647	1	US-08-393-246-7	Sequence 7, Appli
C 32	28.6	8.1	3647	1	US-08-525-058A-7	Sequence 7, Appli
C 33	28.6	8.1	3647	2	US-08-696-731-7	Sequence 7, Appli
C 34	28.6	8.1	3647	4	US-09-042-531-7	Sequence 7, Appli
C 35	28.6	8.1	3647	5	PCT-US91-00899-4	Sequence 4, Appli
C 36	28	8.0	1185	3	US-09-023-339-3	Sequence 3, Appli
C 37	28	8.0	1260	3	US-09-023-173-5	Sequence 5, Appli
C 38	28	8.0	1308	3	US-09-023-173-10	Sequence 10, Appli
C 39	28	8.0	1308	3	US-09-023-339-6	Sequence 6, Appli
C 40	27.6	7.9	1456	4	US-09-056-285A-6	Sequence 6, Appli
C 41	27.6	7.9	1473	4	US-09-056-285A-9	Sequence 9, Appli
C 42	27.6	7.9	1605	4	US-09-149-476-187	Sequence 187, App
C 43	27.6	7.9	1871	1	US-08-416-478A-1	Sequence 1, Appli
C 44	27.6	7.9	1871	2	US-08-474-988B-1	Sequence 1, Appli
C 45	27.6	7.9	1871	2	US-08-394-442B-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-334-220-3
; Sequence 3, Application US/09334220
; Patent No. 6323177
; GENERAL INFORMATION:
; APPLICANT: St. Judes Children's Research Hospital
; APPLICANT: Curran, Thomas
; APPLICANT: D'Arcangelo, Gabriella
; TITLE OF INVENTION: INTERACTION OF REELIN WITH VERY LOW
; TITLE OF INVENTION: DENSITY LIPOPROTEIN (VLDL) RECEPTOR FOR SCREENING AND
; TITLE OF INVENTION: THERAPIES
; FILE REFERENCE: 2427/06704
; CURRENT APPLICATION NUMBER: US/09/334,220
; CURRENT FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 11673
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-334-220-3

Query Match 100.0%; Score 351; DB 4; Length 11673;
Best Local Similarity 100.0%; Pred. No. 1.7e-108;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GAGCAGTGTGGCACCACCATCATGCGCAATGCTGTACCTTCTGTGAGCCGTACGGCCCT	60
Db	970	GAGCAGTGTGGCACCACCATCATGCGCAATGCTGTACCTTCTGTGAGCCGTACGGCCCT	1029
Qy	61	CGAGAGCTGACCCACCATGCTGAACACACACAGCATCTGCTCCCTCCAGTTTTCATT	120
Db	1030	CGAGAGCTGACCCACCATGCTGAACACACACAGCATCTGCTCCCTCCAGTTTTCATT	1089
Qy	121	GGGTGAGGATCATGTCGATTTAGTTACTTCTGACCCAGCATCTGTGTATACGCCAAG	180
Db	1090	GGGTGAGGATCATGTCGATTTAGTTACTTCTGACCCAGCATCTGTGTATACGCCAAG	1149
Qy	181	AACATACCGCTGATTTGATTCAGTTCAGTGAAGAAATAGAGCCCTTCCAATGTGAGACA	240
Db	1150	AACATACCGCTGATTTGATTCAGTTCAGTGAAGAAATAGAGCCCTTCCAATGTGAGACA	1209
Qy	241	GTATTCACATCTCTTACCTCCCGAGGAGCAAGCCAAAGGGAGAGCGTGCAGTTCCAGTGG	300
Db	1210	GTATTCACATCTCTTACCTCCCGAGGAGCAAGCCAAAGGGAGAGCGTGCAGTTCCAGTGG	1269
Qy	301	AAACAGGACACGCTCGAGTGGGTGAGTGATGAGGCCCTGTGGGCCCTG	351
Db	1270	AAACAGGACACGCTCGAGTGGGTGAGTGATGAGGCCCTGTGGGCCCTG	1320

RESULT 2
US-09-334-220-4
; Sequence 4, Application US/09334220
; Patent No. 6323177
; GENERAL INFORMATION:
; APPLICANT: St. Jude Children's Research Hospital
; APPLICANT: Curran, Thomas
; APPLICANT: D'Arcangelo, Gabriella
; TITLE OF INVENTION: INTERACTION OF REELIN WITH VERY LOW
; TITLE OF INVENTION: DENSITY LIPOPROTEIN (VLDL) RECEPTOR FOR SCREENING AND
; TITLE OF INVENTION: THERAPIES
; FILE REFERENCE: 2427/07704
; CURRENT APPLICATION NUMBER: US/09/334,220
; CURRENT FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 11580
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-334-220-4

Query Match 75.1%; Score 263.6; DB 4; Length 11580;
Best Local Similarity 84.6%; Pred. No. 6.2e-79;
Matches 296; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
Qy 1 GAGCAGTGTGGCAGCATCATGCGCAATGCTGTACCTTCTGTGAGCGGTACGGCCT 60
Db 860 GAACAGTGTGGCGGATATGATGCGCAATGCGCTCCTTCTGTGAACCATATGGCCCA 919
Qy 61 CGAGAGTGTACACACATGCTGCAACACAAACAGCATGTGCTCCAGTTTCCATT 120
Db 920 CGAGAACTGATTACACAGGCTTAAACACAAACAGCTTCTGCTCCCAATTTCCATT 979
Qy 121 GGTTCAGATCATGCTGATTTAGTTACTCTGACCCAGCATCATGTGTATACGCCAAG 180
Db 980 GGGTCAGTGTTCATGCTGCTTTAGTTATTCAGACCCAGCATCATGCTGTATATATGCGCAAG 1039
Qy 181 AACATACCGCTGATGATTGATTCAGCTGGAGAAATATAGAGCCCTTCCAAATGTGAGCACA 240
Db 1040 AATACTCTGCGGACTGGATTCAGTACAGAGAAATATAGAGCCCTTCCAAATGTGAGCACA 1099
Qy 241 GTCATCCACATCTGTACCTCCCGAGAGAGCCAAAGAGAGCGTGCAGTTCCAGTGG 300
Db 1100 ATCATCCATATCTCTACCTTCTCTGAGGACGCCAAAGGGGAGATGTCCAATTTTCAGTGG 1159
Qy 301 AACAGGACAGCTGCGAGTGGGTGAGTGTATGAGGCTGCTGGGCTCT 350
Db 1160 AAGCAGGAAATCTTCGTGTAGGTGAAGTGTATGAGGCTGCTGGGCTCT 1209

RESULT 3
US-07-718-575-13
; Sequence 13, Application US/07718575
; Patent No. 5202257
; GENERAL INFORMATION:
; APPLICANT: Heinemann Ph.D., Stephen F.
; APPLICANT: Boulter Ph.D., James R.
; APPLICANT: Hollmann Ph.D., Michael NMN
; APPLICANT: Bettler Ph.D., Bernhard NMN
; APPLICANT: Jensen Ph.D., Jan E.
; TITLE OF INVENTION: GLUTAMATE RECEPTOR COMPOSITIONS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 So. Flower St., Suite 2000
; CITY: Los Angeles
; STATE: California
; COUNTRY: United States
; ZIP: 90071-2921
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/718,575
; FILING DATE: 19910813
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Reiter Ph.D., Stephen E.
; REGISTRATION NUMBER: 31192
; REFERENCE/DOCKET NUMBER: P31 8962
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; TELEX: 9103330318
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3344 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; CELL TYPE:
; CELL LINE:
; IMMEDIATE SOURCE:
; LIBRARY: GluR7
; CLONE: GluR7
; NAME/KEY: CDS
; LOCATION: 1..2766
US-07-718-575-13
Query Match 10.1%; Score 35.6; DB 1; Length 3344;
Best Local Similarity 53.6%; Pred. No. 0.042;
Matches 74; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
Qy 214 ATTAGAGCCCTTCCAAATGTGAGCAGCATCATCCACATCCTGTACCTCCCGAGGAAGCC 273
Db 2317 ATGGCTTCCCTACAGGACAAATCACCATCCCATTTCTGACGTCGAGGAGGAGC 2376
Qy 274 AAGGGGAGAGCGTGCAGTTCCAGTGGAAACAGGACGCTGCGAGTGGGTGAGGTGTAT 333
Db 2377 AAGCTGCACATCATGAAGAGAGTGTGGCGAGGACGCGGTGCCCGGAGGAGAAC 2436
Qy 334 GAGCCTCTGGGCCCTG 351
Db 2437 AAGGAGCCAGCCACTG 2454
RESULT 4
US-08-481-206-13
; Sequence 13, Application US/08481206
; Patent No. 5739291
; GENERAL INFORMATION:
; APPLICANT: Heinemann Ph.D., Stephen F.
; APPLICANT: Boulter Ph.D., James R.
; APPLICANT: Hollmann Ph.D., Michael NMN
; APPLICANT: Bettler Ph.D., Bernhard NMN
; APPLICANT: Jensen Ph.D., Jan E.
; TITLE OF INVENTION: GLUTAMATE RECEPTOR COMPOSITIONS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 So. Flower St., Suite 2000
; CITY: Los Angeles
; STATE: California
; COUNTRY: United States
; ZIP: 90071-2921
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk


```

STATE: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/377,503
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/945,210
FILING DATE: 17-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 1677/192
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3619 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 16..108
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 109..2772
FEATURE:
NAME/KEY: CDS
LOCATION: 16..2772
US-08-377-503-1

```

[illegible]

```

RESULT 7
US-08-178-019-1
; Sequence 1, Application US/08178019
; Patent No. 6333161
; GENERAL INFORMATION:
; APPLICANT: Kamboj, Rajender
; APPLICANT: Elliot, Candace E.
; APPLICANT: Nutt, Stephen L.
; TITLE OF INVENTION: Kainate-Binding, Human CNS Receptors of
; TITLE OF INVENTION: the EAA5 Family
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: Suite 500, 3000 K Street
; CITY: N.W.

```

33

STATE: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/178,019
FILING DATE: 06-JAN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16777/227
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3619 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 16..108
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 109..2772
FEATURE:
NAME/KEY: CDS
LOCATION: 16..2772
US-08-178-019-1

	Query Match	10.1%	Score 35.6;	DB 4;	Length 3619;
	Best Local Similarity	53.6%;	Pred. No.	0.043;	Mismatches
	Matches	74;	Conservative	0;	Mismatches 64; Indels 0; Gaps 0;
Qy	214	ATTAGAGCCCTTCCAATGTGAGCACAGTCATCACAATCGTGACCTCCCCGAGGAAGCC	273		
Dd	2326	ATGGGCTCCCACATACCGGGACAAGATCACCATGCCCATCTCGACTTCAGAGGAGGAGAC	2385		
Qy	274	AAAGGGGAGAGCGTGCAGTTCCAGTGTGCAAAACAGGACAGCCCTGCCAGTGGGTGAGGTGTAT	333		
Dd	2386	AAGCTGCATATCATGAGGAGAAGTGTGTGGGGGGCAGCGGGTGTCTTAGGAGGAGAAAAC	2445		
Qy	334	GAGCCCTGCTGGGCCCCTG	351		
Dd	2446	AAAGAGGCCAGTGCCCTG	2463		

```

RESULT 8
US-09-361-631-4
; Sequence 4, Application US/09361631
; Patent No. 6383778
; GENERAL INFORMATION:
; APPLICANT: Zuker, Jon Elliot
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Lindemeier, Juergen
; TITLE OF INVENTION: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor
; TITLE OF INVENTION: Involved in Sensory Transduction
; FILE REFERENCE: 02307E-08872005
; CURRENT APPLICATION NUMBER: US/09/361,631
; CURRENT FILING DATE: 1999-07-27
; EARLIER APPLICATION NUMBER: US 60/095,464
; EARLIER FILING DATE: 1998-07-28
; EARLIER APPLICATION NUMBER: US 60/112,747
; EARLIER FILING DATE: 1998-12-17

```

33

OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/938, 291A
FILING DATE: September 26, 1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/027,337
FILING DATE: October 11, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 228/172
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 5020 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-938-291A-3

Query Match 8.4%; Score 29.6; DB 3; Length 5020;
Best Local Similarity 56.0%; Pred. No. 5.4;
Matches 56; Conservative 0; Mismatches 44; Indels 0; Gaps 0;
QY 238 ACAGTCATCCATCCCTGACCTCCCGAGGAGCCAAAGGGAGAGCGGTGCAGTTCAG 297
DB 984 ACAGTGTCCACATCTCCCAAGTGTCTCGGTCTGTCACAGCGGGAGCGAGTCTCTCCG 1043
QY 298 TGAACAGGACGACCGCTCGAGTGGGTGAGTGATGAGG 337
DB 1044 CCACAGCATCTCAGAGTGGAGATGCAGATGTGCCAGG 1083

RESULT 12
US-09-488-671-3
Sequence 3, Application US/09488671A
Patent No. 6187545
GENERAL INFORMATION:
APPLICANT: Robert McKay
APPLICANT: Madeline M. Butler
APPLICANT: Jacqueline Wyatt
APPLICANT: Lex M. Cowser
TITLE OF INVENTION: ANTISENSE MODULATION OF PEPCK-CYTOSOLIC EXPRESSION
FILE REFERENCE: RTS-0123
CURRENT APPLICATION NUMBER: US/09/488, 671A
CURRENT FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 177
SEQ ID NO 3
LENGTH: 2657
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (122)...(1990)
US-09-488-671-3

Query Match 8.4%; Score 29.4; DB 4; Length 2657;
Best Local Similarity 55.3%; Pred. No. 4.6;
Matches 57; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 243 CATCCACATCTCTACTCTCCCGAGGAGCCAAAGGGAGAGCGGTTCAGTTCCAGTTGGA 302
DB 247 CATCCACATCTGTGACGGCTCTCAGGAGGAGAAATGGCGGCTTCTGGGCCAGATGGAGGA 306
QY 303 ACAGGACAGCCTCGGAGTGGGTGTATGAGGCTGTCTGG 345
DB 307 AGAGGCGATCCTCAGCGCGTGAAGAAGTATGACAACTGCTGG 349

RESULT 13
US-09-641-806-3
Sequence 3, Application US/09641806
Patent No. 6395527
GENERAL INFORMATION:
APPLICANT: Beraud, Christophe
APPLICANT: Freedman, Richard
TITLE OF INVENTION: No. 6395527el motor proteins and methods for
TITLE OF INVENTION: their use
FILE REFERENCE: 1034
CURRENT APPLICATION NUMBER: US/09/641,806
CURRENT FILING DATE: 2000-08-17
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 1011
TYPE: DNA
ORGANISM: Human
US-09-641-806-3

Query Match 8.3%; Score 29; DB 4; Length 1011;
Best Local Similarity 54.1%; Pred. No. 3.9;
Matches 59; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 207 GGAGAAATATAGAGCCCTTCCAAATGTGAGCAGATCATCCACATCTGTACCTCCCGCA 266
DB 411 GGAGGAGTTCGAGACCTGCTCGAGGTGGGCACTGCCAGCGGTACATCCAGCTCCGGGA 470
QY 267 GGAAGCCAAAGGGAGAGCGGTGCAGTTCAGTTCGAAACAGACAGCGCTG 315
DB 471 AGATGAGCGCGGAATGTTGTGTGTGCGGGGTGAAGGAGGTGCAGCTG 519

RESULT 14
US-09-641-806-1
Sequence 1, Application US/09641806
Patent No. 6395527
GENERAL INFORMATION:
APPLICANT: Beraud, Christophe
APPLICANT: Freedman, Richard
TITLE OF INVENTION: No. 6395527el motor proteins and methods for
TITLE OF INVENTION: their use
FILE REFERENCE: 1034
CURRENT APPLICATION NUMBER: US/09/641,806
CURRENT FILING DATE: 2000-08-17
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1026
TYPE: DNA
ORGANISM: Human
US-09-641-806-1

Query Match 8.3%; Score 29; DB 4; Length 1026;
Best Local Similarity 54.1%; Pred. No. 4;
Matches 59; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 207 GGAGAAATATAGAGCCCTTCCAAATGTGAGCAGATCATCCACATCTGTACCTCCCGCA 266
DB 414 GGAGGAGTTCGAGACCTGCTCGAGGTGGGCACTGCCAGCGGTGACATCCAGCTCCGGGA 473
QY 267 GGAAGCCAAAGGGAGAGCGGTGCAGTTCAGTTCGAAACAGACAGCGCTG 315
DB 474 AGATGAGCGCGGAATGTTGTGTGTGCGGGGTGAAGGAGGTGCAGCTG 522

RESULT 15
US-08-336-408B-3/c
Sequence 3, Application US/08336408B
Patent No. 5723329
GENERAL INFORMATION:
APPLICANT: EVANS, RONALD M.


```

; APPLICANT: MANGELSDORF, DAVID J.
; TITLE OF INVENTION: RETINOID RECEPTOR COMPOSITIONS AND
; TITLE OF INVENTION: METHODS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PRETTY, SCHROEDER, BRUEGEMANN & CLARK
; STREET: 444 SOUTH FLOWER STREET, SUITE 2000
; CITY: LOS ANGELES
; STATE: CA
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,408B
; FILING DATE: 08-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/933,453
; FILING DATE: 21-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US91/00399
; FILING DATE: 22-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/478,071
; FILING DATE: 09-FEB-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: REITER, STEPHEN E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: P41 9851
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-546-1995
; TELEFAX: 619-546-9392
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2095 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 178..1581
; US-08-336-408B-3

```

```

Query Match      8.3%; Score 29; DB 1; Length 2095;
Best Local Similarity 53.0%; Pred. No. 5.6;
Matches 62; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 14 CCATCATGATGCGATGCTGTACCTTCTGTAGCGGTACGGCCCTCGAGAGCTGACCA 73
   ||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1438 CAAACCTGCCCGCTGCTAGGGTACTTGTGTTGCGAGTACGCTTCTAGTGACGCATACA 1379
   ||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | |

QY 74 CCACATCCCTGAACACAGAGCATCTCTCCATTTTCCATTTGGGTCCAGGAT 130
   ||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1378 CCTTCCTCCCTCAACGCTCCACCTCAGAGGGTTTGAGAGCCCTTAGAGTCAGGGT 1322
   ||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: November 6, 2002, 15:28:27
Job time : 81 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 6, 2002, 11:43:51 : Search time 300 Seconds
(without alignments)
2634.838 Million cell updates/sec

Title: US-09-897-438B-1
Perfect score: 351
Sequence: 1 gacgagtgtgaccatcat.....atgagcctctggccctg 351

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	351	100.0	351	24	Mouse reelin prote
2	351	100.0	2745	24	Mus musculus trunc
3	351	100.0	11673	24	Mouse reelin encod
4	351	100.0	11673	24	Mouse ischaemic co
5	263.6	75.1	11580	24	Gene #3437 used to
6	263.6	75.1	11580	24	Human reelin encod
7	263.6	75.1	11632	23	DNA encoding novel
8	195.8	55.8	2025	24	Xenopus laevis int
9	195.8	55.8	2274	24	Xenopus laevis tru

c	10	53.8	15.3	413	22	AAK07720	Human brain expres
c	11	53.8	15.3	413	22	AAK33553	Human bone marrow
c	12	53.8	15.3	413	24	ABS08422	Human genome-deriv
c	13	46.2	13.2	87	22	AAK20384	Human brain expres
c	14	46.2	13.2	87	22	AAK46490	Human bone marrow
c	15	46.2	13.2	87	24	ABS20836	Human genome-deriv
c	16	36.4	10.4	5994	23	ABL10961	Drosophila melanog
c	17	36.4	10.4	8767	23	ABL10960	Drosophila melanog
c	18	35.6	10.1	3619	15	AAQ44685	Excitatory amino a
c	19	32	9.1	1473	24	ABL40040	Synthetic Env gp12
c	20	32	9.1	1986	24	ABL40041	Synthetic Env gp14
c	21	32	9.1	1986	24	ABL40042	Synthetic Env gp14
c	22	32	9.1	2052	24	ABL40050	Synthetic Env gp14
c	23	32	9.1	2073	24	ABL40051	Synthetic Env gp14
c	24	32	9.1	2073	24	ABL40052	Synthetic gp140 po
c	25	32	9.1	2397	24	ABL40043	Synthetic Env gp16
c	26	32	9.1	2529	24	ABL40044	Synthetic Env gp16
c	27	32	9.1	2529	24	ABL40045	Synthetic Env gp16
c	28	32	9.1	2610	24	ABL39970	Synthetic Env gp16
c	29	32	9.1	2613	24	ABL40046	Synthetic Env gp16
c	30	32	9.1	2616	24	ABL40047	Synthetic Env gp16
c	31	32	9.1	2616	24	ABL40048	Synthetic Env gp16
c	32	31.8	9.1	1759	23	ABL01883	Drosophila melanog
c	33	31.8	9.1	4338	23	ABL01882	Drosophila melanog
c	34	31.6	9.0	5332	23	ABL03708	Drosophila melanog
c	35	31.6	9.0	9496	22	AAI02854	Human reproductive
c	36	31.6	9.0	9496	23	ABL97329	Human testicular a
c	37	31.4	8.9	670	24	AA562851	Human G-protein-co
c	38	31.4	8.9	1550	22	AAI71602	Human G-protein co
c	39	31.4	8.9	1604	24	AA36738	Human G-protein co
c	40	31.4	8.9	1860	24	AA562852	Human G-protein-co
c	41	31.4	8.9	2177	24	AA24965	Human G-protein co
c	42	31.4	8.9	5521	22	AAI71603	Human G-protein co
c	43	31.2	8.9	417	22	AAI29056	Colon tumour relat
c	44	31.2	8.9	2203	22	AAK51633	Human polynucleoti
c	45	31.2	8.9	2532	21	AA258964	Mouse GPCR-B4 poly

ALIGNMENTS

RESULT 1

ABL40165	ID	ABL40165	standard; DNA; 351 BP.
XX	AC	ABL40165;	
XX	DT	21-MAY-2002	(first entry)
XX	DE	Mouse reelin protein CR-50	epitope region encoding DNA SEQ ID NO:1.
XX	KW	Mouse; reelin protein CR-50	epitope region; elucidation; neuron;
XX	KW	cerebral disturbance; reelin	protein; neuroprotective; gene; ds.
XX	OS	Mus musculus.	
XX	FH	Key	Location/Qualifiers
XX	FT	CDS	1..351
XX	FT		/*tag= a
XX	FT		/partial
XX	FT		/product= "reelin protein CR-50 epitope region"
XX	FT		/note= "no start or stop codons given"
XX	FN	JP2002017361-A.	
XX	PD	22-JAN-2002.	
XX	PF	04-JUL-2000; 2000JP-0202801.	
XX	PR	04-JUL-2000; 2000JP-0202801.	
XX	PA	(RIKA) RIKAGAKU KENKYUSHO.	

DR WPI; 2002-221707/28.
XX P-PSDB; ABB06244.
PT Reelin protein CR-50 epitope region, useful for diagnosis and treatment
PT of cerebral disturbance
XX
XX Claim 5; Page 11; 16pp; Japanese.
XX
CC The present sequence encodes the mouse reelin protein CR-50 epitope
CC region, which contains the CR-50 antibody recognition site and is free
CC from F-spondin domains and repetitive sites. Also described are: (1) an
CC expression vector comprising a polynucleotide encoding a reelin protein
CC epitope region; (2) host cells with transfected the expression vector;
CC (3) polypeptides prepared by culture of the host cells; and (4)
CC polynucleotides comprising the 351 base sequence given in ABL40165 which
CC encodes the 117 amino acid sequence and/or treatment of diseases caused by
CC the polynucleotide for diagnosis and/or treatment of diseases caused by
CC abnormal positioning of neural cells, and stimulation of association of
CC reelin protein. The mouse reelin protein CR-50 epitope region has
CC neuroprotective activity, and can be used in the diagnosis and treatment
CC of cerebral disturbance due to an abnormal reelin gene and positioning
CC of neurons.
XX
XX Sequence 351 BP; 86 A; 98 C; 92 G; 75 T; 0 other;

Query Match 100.0%; Score 351; DB 24; Length 351;
Best Local Similarity 100.0%; Pred. No. 3.7e-103;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GAGCAGTGTGCACCATCATCGTGGCAATGCTGTCACTTCTGTGAGCCGTAGGCGCT 60
Db 1 GAGCAGTGTGCACCATCATCGTGGCAATGCTGTCACTTCTGTGAGCCGTAGGCGCT 60
Qy 61 CGAGAGTGTGCACCATCATCGTGGCAATGCTGTCACTTCTGTGAGCCGTAGGCGCT 120
Db 61 CGAGAGTGTGCACCATCATCGTGGCAATGCTGTCACTTCTGTGAGCCGTAGGCGCT 120
Qy 121 GGGTCAGGATCATGTCGATTAGTTACTTCTGACCCAGCATCTGTCTATAGCCCAAG 180
Db 121 GGGTCAGGATCATGTCGATTAGTTACTTCTGACCCAGCATCTGTCTATAGCCCAAG 180
Qy 181 AACAAATACCGTGTGATTGATTCAGTGGAGAAATAGAGCCCTTCCAATGTGAGCACA 240
Db 181 AACAAATACCGTGTGATTGATTCAGTGGAGAAATAGAGCCCTTCCAATGTGAGCACA 240
Qy 241 GTCATCCACATCTGTACTCTCCGAGGAGCCAAAGAGGAGCGTGCAGTTCCAGTGG 300
Db 241 GTCATCCACATCTGTACTCTCCGAGGAGCCAAAGAGGAGCGTGCAGTTCCAGTGG 300
Qy 301 AAACAGGACAGCCGTGGGTGGGTGAGTGTATGAGGCGCTGCTGGGCGCTG 351
Db 301 AAACAGGACAGCCGTGGGTGGGTGAGTGTATGAGGCGCTGCTGGGCGCTG 351

RESULT 2
AAD22754
ID AAD22754 standard; cDNA; 2745 BP.
XX
XX AAD22754;
AC
AC
DT 26-FEB-2002 (first entry)
XX
XX Mus musculus truncated reelin cDNA.
DE
XX
XX Mouse; reelin; F-spondin domain; CR-50 epitope; gene therapy; agyria;
KW polymicrogyria; ectopic gray matter; ss.
XX
XX
OS Mus musculus.
XX
XX Key Location/Qualifiers
FT 283..2052
FT CDS /*tag= a
FT /*product= "Mouse truncated reelin protein"

FT sig_peptide 283..363
FT /*tag= b
FT mat_peptide 364..2049
FT /*tag= c
FT /*product= "Mature truncated reelin protein"
FT misc_feature 284..849
FT /*tag= d
FT /*note= "Encodes F-spondin domain"
FT misc_feature 970..1320
FT /*tag= e
FT /*note= "Encodes CR-50 epitope region"
XX
XX EPI149844-A2.
PN 31-OCT-2001.
XX
XX 11-APR-2001; 2001EP-0303411.
PD
XX
XX 11-APR-2000; 2000JP-0109954.
PR
XX
XX (RIKE) RIKEN KK.
PA
XX
XX Mikoshiba K, Tabata H, Nakajima K;
XX
XX WPI; 2002-019320/03.
DR P-PSDB; AAE13606.
XX
XX Novel truncated Reelin protein containing F-spondin domain and CR-50
PT recognition site of Reelin protein, but not having Reelin repeat site,
PT useful to treat diseases including agyria due to abnormal neuron
PT alignment
XX
XX Claim 10; Page 20-26; 47pp; English.
XX
XX The invention relates to a truncated Reelin protein comprising a
CC F-spondin domain and a CR-50 recognition site but no reelin protein
CC recognition site. Reelin is an essential molecule in developing a normal
CC laminated structure of cerebrum. The truncated reelin protein and its
CC DNA are useful for treating diseases including agyria, polymicrogyria,
CC and ectopic gray matter due to abnormal neuronal alignment. Truncated
CC reelin protein DNA is useful in gene therapy. The present sequence is
CC a cDNA encoding Mus musculus truncated reelin protein.
XX
XX Sequence 2745 BP; 661 A; 716 C; 714 G; 654 T; 0 other;
SQ
Query Match 100.0%; Score 351; DB 24; Length 2745;
Best Local Similarity 100.0%; Pred. No. 9e-103;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GAGCAGTGTGCACCATCATCGTGGCAATGCTGTCACTTCTGTGAGCCGTAGGCGCT 60
Db 970 GAGCAGTGTGCACCATCATCGTGGCAATGCTGTCACTTCTGTGAGCCGTAGGCGCT 1029
Qy 61 CGAGAGTGTGCACCATCATCGTGGCAATGCTGTCACTTCTGTGAGCCGTAGGCGCT 120
Db 1030 CGAGAGTGTGCACCATCATCGTGGCAATGCTGTCACTTCTGTGAGCCGTAGGCGCT 1089
Qy 121 GGGTCAGGATCATGTCGATTAGTTACTTCTGACCCAGCATCTGTCTATAGCCCAAG 180
Db 1090 GGGTCAGGATCATGTCGATTAGTTACTTCTGACCCAGCATCTGTCTATAGCCCAAG 1149
Qy 181 AACAAATACCGTGTGATTGATTCAGTGGAGAAATAGAGCCCTTCCAATGTGAGCACA 240
Db 1150 AACAAATACCGTGTGATTGATTCAGTGGAGAAATAGAGCCCTTCCAATGTGAGCACA 1209
Qy 241 GTCATCCACATCTGTACTCTCCGAGGAGCCAAAGAGGAGCGTGCAGTTCCAGTGG 300
Db 1210 GTCATCCACATCTGTACTCTCCGAGGAGCCAAAGAGGAGCGTGCAGTTCCAGTGG 1269
Qy 301 AAACAGGACAGCCGTGGGTGGGTGAGTGTATGAGGCGCTGCTGGGCGCTG 351
Db 1270 AAACAGGACAGCCGTGGGTGGGTGAGTGTATGAGGCGCTGCTGGGCGCTG 1320

Qy	1	GAGCAGCTGTGGCACCACATCATGCAATGCTGTACCTTCTGTGAGCCGTACGCGCCT	60
Db	970	GAGCAGCTGTGGCACCACATCATGCAATGCTGTACCTTCTGTGAGCCGTACGCGCCT	1029
Qy	61	CGAGAGCTGACCACCACCATGCTGAAACAACAACAGACATCTGTCTCCAGTTTTCATT	120
Db	1030	CGAGAGCTGACCACCACCATGCTGAAACAACAACAGACATCTGTCTCCAGTTTTCATT	1089
Qy	121	GGGTCAGGATCATGTCGATTTAGTTACTCTGACCCAGCATCACTGTGTCTATAGCCCAAG	180
Db	1090	GGGTCAGGATCATGTCGATTTAGTTACTCTGACCCAGCATCACTGTGTCTATAGCCCAAG	1149
Qy	181	AAACAATACCGCTGATTTGGATTTCAGCTGGAGAAAATTAGAGCCCTTCCAATTGTCAGCACA	240
Db	1150	AAACAATACCGCTGATTTGGATTTCAGCTGGAGAAAATTAGAGCCCTTCCAATTGTCAGCACA	1209
Qy	241	GTCAATCCATCTCTGTACCTCCCGAGGAAGCAAGGGAGAGCGTGCAGTTCCAGTGG	300
Db	1210	GTCAATCCATCTCTGTACCTCCCGAGGAAGCAAGGGAGAGCGTGCAGTTCCAGTGG	1269
Qy	301	AAACAGGACAGCGTCGAGTGGGTGATGAGGCGCTGCTGGGCGCTG	351
Db	1270	AAACAGGACAGCGTCGAGTGGGTGATGAGGCGCTGCTGGGCGCTG	1320
XX	RESULT 4		
XX	ABI99284		
XX	ID	ABI99284 standard; cdna; 11673 BP.	
XX	ABI99284;		
XX	07-MAR-2002	(first entry)	
XX	Mouse	ischaemic condition related cdna sequence SEQ ID NO:128.	
XX	Mouse;	ischaemia; compressive ischaemia; occlusive ischaemia;	
XX	vasospastic ischaemia;	ischaemic condition; ischaemic disease; ss.	
XX	Mus musculus.		
XX	WO2001188188-A2.		
XX	22-NOV-2001.		
XX	18-MAY-2001;	2001WO-JP04192.	
XX	18-MAY-2000;	2000JP-0145977.	
XX	(UYN1-) UNIV NIHON SCHOOL JURIDICAL PERSON.		
XX	Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;		
XX	WPI: 2002-034733/04.		
XX	P-PSDB; AB57065.		
XX	Examining the ischemic condition (e.g. occlusive ischemia) by measuring		
XX	expression levels of particular genes defined in the specification or		
XX	by determining the expression profile of a gene group comprising these		
XX	genes -		
XX	Claim 2; Page 362-385; 2690pp; English.		
XX	The present invention describes a method for examining ischaemic		
XX	conditions, comprising measuring the expression levels of particular		
XX	genes (1) in a test sample or determining the expression profile of a		
XX	gene group in the sample comprising genes selected from (1). The method		
XX	is useful for examining the ischaemic condition (e.g. compressive		
XX	ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring		
XX	expression levels of particular genes (ABI99202 to ABI99912, encoding		
XX	the protein sequences in AB57020 to AB57374) or by determining the		
XX	expression profile of a gene group comprising these genes. The		
XX	expression levels or expression profiles produced by these genes are		
XX	used as an indicator when screening for ischaemic condition-improving		

CC drugs or therapeutics for ischaemic diseases. AB199913 and AB199914
CC represent PCR primers for a mouse ischaemic condition related sequence,
CC which are used in the exemplification of the present invention.
XX
SQ Sequence 11673 BP; 2831 A; 2985 C; 2985 G; 2872 T; 0 other;

Query Match 100.0%; Score 351; DB 24; Length 11673;
Best Local Similarity 100.0%; Pred. No. 1.7e-102;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GAGCAGTGTGCGACCATCATGCGCAATGCTGTCACTTCTGTGAGCGGTACGGCCCT 60
Dy 970 GAGCAGTGTGCGACCATCATGCGCAATGCTGTCACTTCTGTGAGCGGTACGGCCCT 1029
Qy 61 CGAGAGTGTGCGACCATCATGCGCAATGCTGTCACTTCTGTGAGCGGTACGGCCCT 120
Dy 1030 CGAGAGTGTGCGACCATCATGCGCAATGCTGTCACTTCTGTGAGCGGTACGGCCCT 1089
Qy 121 GGGTCAGATCATGCTGATTTAGTTACTTCTGACCTTCTGTGAGCGGTACGGCCCT 180
Dy 1090 GGGTCAGATCATGCTGATTTAGTTACTTCTGACCTTCTGTGAGCGGTACGGCCCT 1149
Qy 181 AACATACCGCTGATTTGATTTAGTTACTTCTGACCTTCTGTGAGCGGTACGGCCCT 240
Dy 1150 AACATACCGCTGATTTGATTTAGTTACTTCTGACCTTCTGTGAGCGGTACGGCCCT 1209
Qy 241 GTCATCCACATCCTGTACCTCCCGAGAGAGCCAAAGGGGAGAGCGTCCAGTTCACGTG 300
Dy 1210 GTCATCCACATCCTGTACCTCCCGAGAGAGCCAAAGGGGAGAGCGTCCAGTTCACGTG 1269
Qy 301 AACAGGACAGCGTGGAGTGGGTGAGGTGTATGAGCGCTGCTGGGCCCTG 351
Dy 1270 AACAGGACAGCGTGGAGTGGGTGAGGTGTATGAGCGCTGCTGGGCCCTG 1320

RESULT 5
ID ABN96939 standard; DNA; 11580 BP.
XX
AC ABN96939;
XX
DT 13-AUG-2002 (first entry)
DE Gene #3437 used to diagnose liver cancer.
XX

Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
metastatic liver tumour; cytostatic; expression profile; disease state;
disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.

XX WO200229103-A2.
XX 11-APR-2002.
XX
XX 02-OCT-2001; 2001WO-US30589.
XX
XX 02-OCT-2000; 2000US-237054P.
XX
XX (GENE-) GENE LOGIC INC.
XX

PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX WPI; 2002-426119/45.
XX

XX Diagnosing and detecting the progression of liver cancer,
PT hepatocellular carcinoma or metastatic liver tumor in a patient,
PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample -
XX

XX Claim 1; SEQ ID NO 3437; 298pp; English.
XX
XX The invention relates to a novel method for diagnosing and detecting the

CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumor in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 11580 BP; 3014 A; 2696 C; 2753 G; 3116 T; 1 other;
Query Match 75.1%; Score 263.6; DB 24; Length 11580;
Best Local Similarity 84.6%; Pred. No. 2.8e-74;
Matches 296; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
Qy 1 GAGCAGTGTGCGACCATCATGCGCAATGCTGTCACTTCTGTGAGCGGTACGGCCCT 60
Dy 860 GAACAGTGTGCGCGCATATGCGATGGCAATGCGTCACTTCTGTGAACCATATGGCCCA 919
Qy 61 CGAGAGTGTGCGACCATCATGCGCTGAACACACACACATCTGTCTCCAGTTTTCATT 120
Dy 920 CGAAGACTGATTAACACAGCGCTTAATACACACACAGCTTCTGTCTCCCAATTTCCATT 979
Qy 121 GGGTCAGATCATGCTGATTTAGTTACTTCTGACCTTCTGTGAGCGGTACGGCCCT 180
Dy 980 GGGTCAGATCATGCTGCTTTAGTTACTTCTGACCTTCTGTGAGCGGTATATGCGCAAG 1039
Qy 181 AACAAATACCGCTGATTTGATTTAGTTACTTCTGACCTTCTGTGAGCGGTACGGCCCT 240
Dy 1040 AATAACTCTCGGACATGATTTAGTTACTTCTGAGAGAAATTAAGAGCCCTTCCAATGTCAGCACA 1099
Qy 241 GTCATCCACATCCTGTACCTCCCGAGAGAGCCAAAGGGGAGAGCGTCCAGTTCACGTG 300
Dy 1100 ATCATCCATATCTCTACCTTCTGTGAGGAGCGCCCAAGGGGAGAGTGTCCAATTTCACTGG 1159
Qy 301 AACAGGACAGCGTGGAGTGGGTGAGGTGTATGAGCGCTGCTGGGCCCT 350
Dy 1160 AAGCAGGAGAAATCTTCTGTAGGTGAAGTGTATGAAGCCCTGCTGGGCCCT 1209

RESULT 6
ID ABA92604 standard; cDNA; 11580 BP.
XX
AC ABA92604;
XX
DT 21-MAR-2002 (first entry)
DE Human reelin encoding cDNA SEQ ID NO:4.
XX

XX Human; reelin; low density lipoprotein receptor; LDLR; neuroprotective;
KW extracellular glycoprotein; nootropic; antilipemic; Alzheimer's disease;
KW neurodegenerative disorder; neuronal regeneration; cognitive function;
KW lipid metabolism disease; memory; developmental disorder; gene; ss.
XX

OS Homo sapiens.
XX
XX Key Location/Qualifiers
FT CDS 176..10558
FT /*tag= a
FT /product= "human reelin"
XX
XX US6323177-B1.
XX
XX PD 27-NOV-2001.
XX
XX 16-JUN-1999; 99US-0334220.


```
Db 980 GGGTCAGGTTTCGCTTGTAGTTATTCAGACCCAGCATCATCGTGTATATGCCAAG 1039
QY 181 AACAAATACCCGTCGATTCAGTCAGTCGAGAGAAATAGAGCCCTTCCAATGTGAGCACA 240
Db 1040 AATAACTCTCGGACTGGATTCAGTCAGTAGAGAAATAGAGCCCTTCCAATGTGAGCACA 1099
QY 241 GTCATCCACATCTGTACCTCCCGAGGAGGAGCCAAAGGGAGGAGCGTGCAGTTCAGTGG 300
Db 1100 ATCATCCATATCTCTACCTTCCTGAGGAGGCCAAAGGGAGGAGATGTCCAATTCAGTGG 1159
QY 301 AACAGGACAGCCGTCGAGTGGGTGAGTGTATCAGGCCCTGCTGGGCCCT 350
Db 1160 AAGCAGGAAATCTTCGTGAGGTGAAGTGTATCAGGCCCTGCTGGGCCCTT 1209

RESULT 8
AAD22779
ID AAD22779 standard; cDNA; 2025 BP.
AC AAD22779;
XX
DT 26-FEB-2002 (first entry)
DE Xenopus laevis intact reelin cDNA.
XX
KW Frog; reelin; F-spondin domain; CR-50 epitope; gene therapy; agyria;
KW polymicrogyria; ectopic gray matter; ss.
XX
OS Xenopus laevis.
XX
FH Key Location/Qualifiers
FT CDS 157..2025
FT /tag= a
FT /product= "Xenopus intact reelin protein"
FT /transl_except= (pos:1606..1608, aa:Cys)
FT /transl_except= (pos:1639..1641, aa:Cys)
FT /note= "CDS does not include stop codon"
FT /partial
XX
PN EP1149844-A2.
XX
PD 31-OCT-2001.
XX
PF 11-APR-2001; 2001EP-0303411.
XX
PR 11-APR-2000; 2000JP-0109954.
XX
PA (RIKE ) RIKEN KK.
XX
PI Mikoshiba K, Tabata H, Nakajima K;
XX
WPI; 2002-019320/03.
DR P-PSDB; AAEL3607.
XX
PT Novel truncated Reelin protein containing F-spondin domain and CR-50
PT recognition site of Reelin protein, but not having Reelin repeat site,
PT useful to treat diseases including agyria due to abnormal neuron
PT alignment -
XX
PS Example 1; Fig 1A; 47pp; English.
XX
CC The invention relates to a truncated Reelin protein comprising a
CC F-spondin domain and a CR-50 recognition site but no reelin protein
CC repeat site. Reelin is an essential molecule in developing a normal
CC laminated structure of cerebrum. The truncated reelin protein and its
CC DNA are useful for treating diseases including agyria, polymicrogyria,
CC and ectopic gray matter due to abnormal neuronal alignment. Truncated
CC reelin protein DNA is useful in gene therapy. The present sequence is
CC a cDNA encoding Xenopus laevis intact reelin protein.
XX
SQ Sequence 2025 BP; 541 A; 419 C; 460 G; 604 T; 1 other;
e. Query Match 55.8%; Score 195.8; DB 24; Length 2025;
```

```
Best Local Similarity 72.4%; Pred. No. 1e-52;
Matches 254; Conservative 0; Mismatches 97; Indels 0; Gaps 0;
QY 1 GAGCAGTGTGGCACCACATCATGCAATGCTGTACCTTCTGTGAGCGGTAGGGCCCT 60
Db 847 GAGCAGTGTGGTGCATTTATGCAATGTTGGGGCAGTCACTTTTGTGATCCGTATGGACCA 906
QY 61 CGAGAGCTGACCCACCATGCCTGACACAAACACAGCATCTGTCTCCAGTTTCCATT 120
Db 907 AGAAGATTGATAACTGTTCAATGAACACAACTACGGCATCTGTTTGCAGTTTTCTATT 966
QY 121 GGTTCAGGATCATGTCGATTTAGTTACTCTGACCCAGCATCACTGTGTCTATAGCCCAAG 180
Db 967 GGGTCAGGATCGTGCAGGTTTCAGCTATTCAGACCCCTGGAATTTGTGTGTATACAAAG 1026
QY 181 AACAAATACCCGTCGATTCAGTCAGTCGAGAGAAATAGAGCCCTTCCAATGTGAGCACA 240
Db 1027 AATAATTCATCTAGTTGGATGCCATTCGAGAGAAATAGTCTCTTCCAATGTAGCACC 1086
QY 241 GTCATCCACATCTGTACCTCCCGAGGAGGAGCCAAAGGGAGGAGCGTGCAGTTCAGTGG 300
Db 1087 ATCATTACACATTATTTACCTACCTACCTGAAAGCTAAAGGAGAAATGTGAAATCCGTTGG 1146
QY 301 AACAGGACAGCCGTCGAGTGGGTGAGTGTATCAGGCCCTGCTGGGCCCTG 351
Db 1147 AGGAGGAGAAACATGCAGGCGAGGTGATGTGTATGAGCCCTGCTGGGCACTG 1197

RESULT 9
AAD22753
ID AAD22753 standard; cDNA; 2274 BP.
XX
AC AAD22753;
XX
DT 26-FEB-2002 (first entry)
DE Xenopus laevis truncated reelin cDNA.
XX
KW Frog; reelin; F-spondin domain; CR-50 epitope; gene therapy; agyria;
KW polymicrogyria; ectopic gray matter; ss.
XX
OS Xenopus laevis.
XX
FH Key Location/Qualifiers
FT CDS 157..1455
FT /tag= a
FT /product= "X. laevis truncated reelin protein"
FT sig_peptide 157..234
FT mat_peptide /tag= b
FT 235..1452
FT /tag= c
FT misc_feature /product= "Mature truncated reelin protein"
FT 241..726
FT /tag= d
FT /note= "Encodes F-spondin domain"
FT 847..1197
FT /tag= e
FT /note= "Encodes CR-50 epitope region"
FT 2254..2259
FT /tag= f
XX
PN EP1149844-A2.
XX
PD 31-OCT-2001.
XX
PF 11-APR-2001; 2001EP-0303411.
XX
PR 11-APR-2000; 2000JP-0109954.
XX
PA (RIKE ) RIKEN KK.
XX
PI Mikoshiba K, Tabata H, Nakajima K;
XX
```



```
DR WPI: 2002-019320/03.
XX P-PSDB; AAE13605.
PT Novel truncated Reelin protein containing F-spondin domain and CR-50
PT recognition site of Reelin protein, but not having Reelin repeat site,
PT useful to treat diseases including agyria due to abnormal neuron
PT alignment
XX
XX Claim 8; Page 11-16; 47pp; English.
XX
XX The invention relates to a truncated Reelin protein comprising a
XX F-spondin domain and a CR-50 recognition site but no reelin protein
XX repeat site. Reelin is an essential molecule in developing a normal
XX laminated structure of cerebrum. The truncated reelin protein and its
XX cDNA are useful for treating diseases including agyria, polymicrogyria,
XX and ectopic gray matter due to abnormal neuronal alignment. Truncated
XX reelin protein DNA is useful in gene therapy. The present sequence is
XX a cDNA encoding Xenopus laevis truncated reelin protein.
XX
XX Sequence 2274 BP; 660 A; 450 C; 469 G; 694 T; 1 other;
XX
XX Query Match 55.8%; Score 195.8; DB 24; Length 2274;
XX Best Local Similarity 72.4%; Pred. No. 1e-52;
XX Matches 254; Conservative 0; Mismatches 97; Indels 0; Gaps 0;
XX
QY 1 GAGCAGTGTGGCACCACATCATGCGCATGCGCAATGCTGTCACTTCTGTGAGCCGTACGGCCCT 60
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 847 GAGCAGTGTGGTGCATATATGCGATGGTGGGCGACGTCACTTTTGTGATCGGTATGGACCA 906
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 CGAGAGCTGACCCACATCGCTGAACACAAACACACATCTGTCTCCAGTTTTCATT 120
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 907 AGAATGATACTGTTCAATGAACACACACTACGCACTGTTTGCAGCTTTCTATT 966
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 GGTGAGATCATGTCGATTAGTTACTGTGACCCAGCATCATGTCATACGCCAAG 180
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 967 GGTGAGATCATGTCGATTAGTTACTGTGACCCAGCATCATGTCATACGCCAAG 1026
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 AACAAACCCGTGATTGATTCAGCTCGAGAAATAGACCCCTTCCAATGTGAGCACA 240
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1027 AATAATTCATCTAGTTGGATGCCATTGGAGAAATAGTGTCTCTCCAATGTTAGCACC 1086
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 GTATCCACATCTCTGTACCTCCCGAGGAGCCAAAGGAGAGCGTGCAGTTTCCAGTGG 300
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1087 ATCATTCACATATTATTACCTACCTCTGAAGCTAAGAGAGAAATGTGAATTCCTGTTG 1146
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 301 AAACAGGACAGCCGTCGAGTGGGTGAGGTGTATGAGGCCCTGCTGGGCCCTG 351
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1147 AGCAGGAGAAACATGCAGGCGAGTGTGTGTATGAAGCCTGCTGGGCACCTG 1197
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
AAK07720/c
ID AAK07720 standard; DNA; 413 BP.
XX
XX AAK07720;
XX
XX 05-NOV-2001 (first entry)
XX
XX Human brain expressed single exon probe SEQ ID NO: 7711.
XX
XX DE
XX KW Human; brain expressed exon; gene expression analysis; probe;
XX microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX epilepsy; cancer; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200157275-A2.
XX
XX XX
XX PD 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00667.
XX
XX 04-FEB-2000; 2000US-0180312.
XX
XX PR

26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains
XX
XX Example 4; SEQ ID NO: 7711; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is one of the probes of the
XX invention.
XX
XX Sequence 413 BP; 134 A; 86 C; 67 G; 126 T; 0 other;
XX
XX Query Match 15.3%; Score 53.8; DB 22; Length 413;
XX Best Local Similarity 79.0%; Pred. No. 3.4e-07;
XX Matches 64; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
XX
QY 1 GAGCAGTGTGGCACCACATCATGCGCATGCGCAATGCTGTCACTTCTGTGAGCCGTACGGCCCT 60
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 CGAGAGCTGACCCACACATGTC 81
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 263 CGAAGAACTGGTAAGTATGTGC 243
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 11
AAK33553/c
ID AAK33553 standard; DNA; 413 BP.
XX
XX AAK33553;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human bone marrow expressed single exon probe SEQ ID NO: 8110.
XX
XX DE
XX KW Human; bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200157276-A2.
XX
XX XX
XX PD 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00668.
XX
XX 04-FEB-2000; 2000US-0180312.
XX
XX 26-MAY-2000; 2000US-0207456.
XX
XX 30-JUN-2000; 2000US-0608408.
XX
XX 03-AUG-2000; 2000US-0632366.
XX
XX 21-SEP-2000; 2000US-0234687.
XX
XX 27-SEP-2000; 2000US-0236359.
XX
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains
XX
XX Example 4; SEQ ID NO: 7711; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is one of the probes of the
XX invention.
XX
XX Sequence 413 BP; 134 A; 86 C; 67 G; 126 T; 0 other;
XX
XX Query Match 15.3%; Score 53.8; DB 22; Length 413;
XX Best Local Similarity 79.0%; Pred. No. 3.4e-07;
XX Matches 64; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
XX
QY 1 GAGCAGTGTGGCACCACATCATGCGCATGCGCAATGCTGTCACTTCTGTGAGCCGTACGGCCCT 60
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 323 GAACAGTGTGGCGGATTTATGCGATGCGCAATGCCGTCACTTCTGTGAACCATATGGCCCA 264
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 CGAGAGCTGACCCACACATGTC 81
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 263 CGAAGAACTGGTAAGTATGTGC 243
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```

XX WPI; 2001-488900/53.
DR
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -
PT
XX Example 4; SEQ ID NO: 8110; 659pp + Sequence Listing; English.
PS
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention.
XX
SQ Sequence 413 BP; 134 A; 86 C; 67 G; 126 T; 0 other;
    Query Match          15.3%; Score 53.8; DB 22; Length 413;
    Best Local Similarity 79.0%; Pred. No. 3.4e-07;
    Matches 64; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 1 GAGCAGTGTGGCACCATCATGCATGTCACCTTCTGTGAGCGGTACGCCCT 60
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 323 CGACAGTGTGGCGGATTATGCATGCGCATGCCGTCACTTCTGTGAACCATATGCCCA 264
QY 61 GGACAGTGACACCACCATGTC 81
   ||||| ||| | | |||
Db .263 CGAAGACTGGTAAGATGTGC 243
RESULT 12
ABS08422/c
ID ABS08422 standard; DNA; 413 BP.
XX AC ABS08422;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe from lung SEQ ID No 8413.
XX
KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US00665.
XX
PR 04-FEB-2000; 2000US-180312P.
XX
PR 26-MAY-2000; 2000US-207456P.
XX
PR 30-JUN-2000; 2000US-060840B.
XX
PR 03-AUG-2000; 2000US-063236G.
XX
PR 21-SEP-2000; 2000US-234687P.
XX
PR 27-SEP-2000; 2000US-236359P.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to

```

PT measure gene expression in human lung samples -

XX Claim 1; SEQ ID NO 8413; 634pp; English.

XX

CC The invention relates to a spatially-addressable set of single exon

CC nucleic acid probes for measuring gene expression in a sample derived

CC from human lung comprising single exon nucleic acid probes having one of

CC 12614 nucleic acid sequences mentioned in the specification, or their

CC complements or the 12387 open reading frames derived from the 12614

CC probes. Also included are a microarray comprising the novel set of

CC probes; the novel set of probes which hybridise at high stringency to a

CC nucleic acid expressed in the human lung; measuring gene expression in a

CC sample derived from human lung, comprising (a) contacting the array with

CC a collection of detectably labeled nucleic acids derived from human lung

CC mRNA; and (b) measuring the label detectably bound to each probe of

CC the array; identifying exons in a eukaryotic genome, comprising

CC (a) algorithmically predicting at least one exon from genomic sequences

CC of the eukaryote; and (b) detecting specific hybridisation of detectably

CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,

CC having a fragment identical to the predicted exon, the probe is included

CC in the above mentioned microarray; assigning exons to a single gene,

CC comprising (a) identifying exons from genomic sequence by the method

CC above and (b) measuring the expression of each of the exons in several

CC tissues and/or cell types using hybridisation to a single exon

CC microarrays having a probe with the exon, where a common pattern of

CC expression of the exons in the tissues and/or cell types indicates that

CC the exons should be assigned to a single gene; a peptide comprising one

CC of 12011 sequences, mentioned in the specification, or encoded by the

CC probes/open reading frames (ORF). The probes are used for gene

CC expression analysis, and for identifying exons in a gene, particularly

CC using human lung derived mRNA and for the study of lung diseases

CC such as asthma, lung cancer, chronic obstructive pulmonary disease

CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary

CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,

CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary

CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,

CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic

CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension

CC and hyaline membrane disease. The present sequence is a single exon

CC probe of the invention.

CC Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic

CC format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences.

XX

XX Sequence 413 BP; 134 A; 86 C; 67 G; 126 T; 0 other;

XX

Query Match 15.3%; Score 53.8; DB 24; Length 413;

Best Local Similarity 79.0%; Pred. No. 3.4e-07;

Matches 64; Conservative 0; Mismatches 17; Indels 0; Gaps 0

Qy 1 GAGCAGTGTGGCACCATTATCGATGCAATGCTGTGTCACCTCTGTGAGCCGTACGGCCCT 60

Db 323 GAACAGTGTGGCGGATATGTCATGCAATGCGTCACCTCTGTGAAACCATATGGCCA 264

Qy 61 CGAGAGCTGTACCACCATGCG 81

||||| ||| | |||

Db 263 CGAGAACTGTGAAGTATGTC 243

||||| ||| | |||

RESULT 13

AAK20384/C

ID AAK20384 standard; DNA; 87 BP.

XX AAK20384;

XX 05-NOV-2001 (first entry)

XX

XX Human brain expressed single exon probe SEQ ID NO: 20375.

XX

XX Human; brain expressed exon; gene expression analysis; probe;

XX microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

XX epilepsy; cancer; ss.

KW

SQ Sequence 87 BP; 21 A; 25 C; 19 G; 22 T; 0 other;

Qy 1 GAGCAGTGTGGCACCATCATGCATGCGCAATGCTGTTCACCTTCTGTGAGCCGTACGGCCC 59
|| ||||||| || ||||||||||| ||||||||||| ||||||| || |||||||
Dd 59 GAACAGTGTGGCGCGATTATGCATGGCAATGCCGTCACTTCTGTGAACCAATATGGCCC 1

Search completed: November 6, 2002, 14:03:41
Job time : 312 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 6, 2002, 13:54:36 ; Search time 2213 Seconds
(without alignments)
2568.737 Million cell updates/sec

Title: US-09-897-438B-1
Perfect score: 351
Sequence: 1 gagcagtggcaccatcat.....atgagccctgctggccctg 351

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```
EST: *
1: em_estba: *
2: em_esthum: *
3: em_estnu: *
4: em_estnu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_hic: *
9: gb_esti: *
10: gb_est2: *
11: gb_hic: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: gb_gss: *
18: em_gss_hum: *
19: em_gss_inv: *
20: em_gss_pln: *
21: em_gss_vrt: *
22: em_gss_fun: *
23: em_gss_mam: *
24: em_gss_mus: *
25: em_gss_other: *
26: em_gss_pro: *
27: em_gss_rod: *
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	203.8	58.1	315	12	BE938667 QV0-TN008
C 2	145	41.3	600	12	BG803882 0243-03 M
C 3	94	26.8	635	10	BB248113 BB248113
C 4	64	18.2	568	12	BF387628 UT-R-CAL-
C 5	39.6	11.3	943	14	BQ647080 AGENCOURT
C 6	39.2	11.2	514	14	BQ955513 QGG10D11.

7	38.6	11.0	1022	17	CNS012TZ
8	36	10.3	547	13	BI803750
C 9	35.6	10.1	993	17	CNS03750
C 10	35.4	10.1	330	10	AW922767
C 11	34.8	9.9	374	17	AZ777615
C 12	34.8	9.9	582	10	AV522139
C 13	34.8	9.9	668	13	BI573885
C 14	34.6	9.9	528	14	BQ038901
C 15	34.6	9.9	917	12	BF135143
C 16	34.2	9.7	507	10	AW920326
C 17	34	9.7	505	14	BM956552
C 18	34	9.7	547	12	BE80801
C 19	33.8	9.6	872	17	CNS03DM3
C 20	33.6	9.6	435	17	AZ558748
C 21	33.4	9.5	563	12	BG420388
C 22	33.4	9.5	1077	17	CNS04XQV
C 23	33.2	9.5	476	12	BF499004
C 24	33.2	9.5	485	17	AZ822150
C 25	33.2	9.5	691	12	BG600888
C 26	33.2	9.5	715	13	BM622337
C 27	33.2	9.5	725	12	BF492454
C 28	33	9.4	332	10	AW401228
C 29	33	9.4	668	17	AQ475926
C 30	32.8	9.3	412	12	BF288465
C 31	32.6	9.3	589	14	BM715558
C 32	32.6	9.3	874	9	AL554368
C 33	32.6	9.3	938	17	CNS03BLG
C 34	32.4	9.2	419	17	AQ237938
C 35	32.4	9.2	431	12	BF603741
C 36	32.4	9.2	465	9	AA835142
C 37	32.4	9.2	701	17	AQ316104
C 38	32.4	9.2	851	17	BH129632
C 39	32.4	9.2	1038	17	CNS05S4W
C 40	32.2	9.2	412	9	AA617544
C 41	32.2	9.2	483	12	BF441217
C 42	32.2	9.2	719	10	BE614332
C 43	32	9.1	455	10	AW566596
C 44	32	9.1	550	13	BI865114
C 45	32	9.1	559	13	BI318806

ALIGNMENTS

RESULT 1
BE938667/C
LOCUS
DEFINITION QV0-TN0084-180800-342-a08 TN0084 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE938667
VERSION BE938667.1 GI:10466224
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W., Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 315)

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

Excerpt

Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2=QV0-TN0084-180
800-342-a086t3-2000-08-186t4-1)
Seq primer: puc 18 forward
High quality sequence stop: 315.
Location/Qualifiers
1. .315

FEATURES

source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="TN0084"
/dev_stage="Adult"
/note="Organ: testis_normal; Vector: puc18; Site:1: Smal;
Site:2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
84 a 64 c 78 g 89 t

BASE COUNT

ORIGIN

Query Match 58.1%; Score 203.8; DB 12; Length 315;
Best Local Similarity 84.5%; Pred. No. 9.2e-50;
Matches 229; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY 80 GCCTTAACACAAACAGCATCTCTCCATGTTTCCATTTGGCTGAGGATCATGTCGAT 139
DB 306 GCCTTAACACAAACAGCATCTCTCCATTTGGCTGAGGATCATGTCGCT 247
QY 140 TTAGTTACTCTGACCCAGCATCATCTGTCTACGCAAGCAATACCGCTGATTGGA 199
DB 246 TTAGTTATTGACCCAGCATCATCTGTATATGCAAGATAACTCTGCGACTGGA 187
QY 200 TTCAGCTGGAGAAATTTAGAGCCCTTCCATGTGACACAGTCATCCATCTGTACC 259
DB 186 TTCAGCTAGAGAAATTTAGAGCCCTTCCATGTGACACAAATCATCTCTTACC 127
QY 260 TCCCGGAGGAAGCCAAAGGGGAGGCTGCGAGTTCCAGTGGAAACAGCAGCCTGCGAG 319
DB 126 TTCTCTGAGGACCCCAAGGGGAGAGATGTCCAAATTCAGTGGAGCAGCAAGATCTTCGTG 67
QY 320 TGGGTGAGGTGTATGAGGCTGCTGGGCCCT 350
DB 66 TAGCTGAAGTGTATGAAGCCTGCTGGGCCCT 36

RESULT 2

BG803882/c

LOCUS

0243-03 Mouse El4.5 retina lambda ZAP II Library Mus musculus cDNA,
mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

Contact: Klein WH

Department of Biochemistry and Molecular Biology

University of Texas M.D. Anderson Cancer Center

Box 117, 1515 Holcombe Blvd., Houston, TX 77030, USA

Tel: 713 792 3646

Fax: 713 790 0329.

FEATURES

source

Location/Qualifiers
1. .600
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="Mouse El4.5 retina lambda ZAP II Library"
/tissue_type="neural retina"
/dev_stage="embryonic day 14.5 post-fertilization"
/note="Vector: pAMP10 (Gibco); Cloned unidirectionally.
Primer: Oligo dT. RNA Isolation: cytoplasmic RNA preps
(Mannatis); Cloning Technique: CUA Cloning (CloneAmp,
Life Technologies); Average insert size: 1.8 Kb;
Insertion site: TAGCTCCACTGAATCTGAGTG--->. Other
information regarding entire library may be found at
http://pga.swmed.edu/Data/Libraries/microarray_cdna_librar
ies.htm."

BASE COUNT 127 a 167 c 152 g 152 t 2 others

ORIGIN

Query Match 41.3%; Score 145; DB 12; Length 600;
Best Local Similarity 71.8%; Pred. No. 3.1e-32;
Matches 252; Conservative 0; Mismatches 2; Indels 97; Gaps 1;
QY 1 GAGCAGTGTGGCACCATCATGTCGCAATGCTGTCACCTCTCTGTGAGCGGTAGGGCCCT 60
DB 596 GAGCAGTGTGGCACCATCATGTCGCAATGCTGTCACCTCTCTGTGAGCGGTAGGGCCCT 537
QY 61 CGAGAGCTGACCCACACATGCTGCAACACAAACAGCATCTGCTCCAGTTTCCATT 120
DB 536 CGAGAGCTGACCCACACATGCTGCAACACAAACAGCATCTGCTCCAGTTTCCATT 477
QY 121 GGCTCAGGATCATGTGCGATTGATTACTTGTGACCCAGCATCACTGTGTATACGCGCAAG 180
DB 476 ----- 477
QY 181 AACAAATACCGCTGATTCGATTCAGCTGGAGAAATTTAGAGCCCTTCCAAATGTGAGCACA 240
DB 476 -----GAGCCCCCTTCCAAATGTGAGCACA 454
QY 241 GTCATCCACATCCTGTACCTCCCGAGGAGCCAAAGGAGAGCGTGCAGTTCCAGTGG 300
DB 453 GTCATCCACATCCTGTACCTCCCGAGGAGCCAAAGGAGAGCGTGCAGTTCCAGTGG 394
QY 301 AAACAGGACAGCCTCGCAGTGGGTGAGTGTATGAGGCGCTGCTGGGCCCTG 351
DB 393 AAACAGGACAGCCTCGCAGTGGGTGAGTGTATGAGGCGCTGCTGGGCCCTG 343

RESULT 3

BB248113

LOCUS

BB248113 RIKEN full-length enriched, 635 bp mRNA linear EST 23-OCT-2001
musculus cDNA clone A730023J04 3', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (2001)

On Jul 6, 2000 this sequence version replaced gi:8940859.

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa
K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
Hayashizaki, Y.

Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
further details.

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

FEATURES

source

```

1. .635
  /organism="Mus musculus"
  /db_xref="taxon:10090"
  /clone_lib="A73023J04"
  /clone_lib="RIKEN full-length enriched, 7 days neonate
  cerebellum"
  /tissue_type="cerebellum"
  /dev_stage="7 days neonate"
  /lab_host="DH10B"
  /notes="Site_1: SalI; Site_2: BamHI; cDNA library was
  prepared and sequenced in Mouse Genome Encyclopedia
  Project of Genome Exploration Research Group in Riken
  Genomic Sciences Center and Genome Science Laboratory in
  RIKEN. Division of Experimental Animal Research in Riken
  contributed to prepare mouse tissues. 1st strand cDNA was
  primed with a primer [5'
  GAGAGAGAGAGATCCAGAGCTCTTTTTTTTTTTVN 3'], cDNA was
  prepared by using trehalase thermo-activated reverse
  transcriptase and subsequently enriched for full-length by
  cap-trapper. cDNA went through one round of normalization
  to Rot = 20.0 and subtraction to Rot = 459.0. Second
  strand cDNA was prepared with the primer adapter of
  sequence [5' GAGAGAGATTCGAGTTAATTAATTAATCCGCCGCC
  3']. cDNA was cleaved with XhoI and BamHI. Vector: a
  modified pBluescript KS(+) after bulk excision from Lambda
  FLC I."
```

BASE COUNT 182 a 122 c 118 g 213 t
ORIGIN

Query Match 26.8%; Score 94; DB 10; Length 635;
Best Local Similarity 100.0%; Pred. No. 4,1e-17;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 29 ATGCTGTACCTTCTGTGAGCCCTACGCCCTCGAGAGCTGACACACACATGCTGAACA 88
|||||
Db 1 ATGCTGTACCTTCTGTGAGCCCTACGCCCTCGAGAGCTGACACACATGCTGAACA 60
|||||

Qy 89 CAACAACAGCATCTGCTCCAGTTTCCATTGG 122
|||||
Db 61 CAACAACAGCATCTGCTCCAGTTTCCATTGG 94
|||||

RESULT 4

BF387628/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

BF387628 568 bp mRNA linear EST 27-NOV-2000
UI-R-CA1-bbs-e-06-0-UI.s1 UI-R-CA1 Rattus norvegicus cDNA clone
UI-R-CA1-bbs-e-06-0-UI 3', mRNA sequence.

BF387628

BF387628.1 GI:11372452

EST.

Norway rat.

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 568)

Bonaldo, M.F., Lennon, G. and Soares, M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

9704477

Contact: Soares, MB

Program for Rat Gene Discovery and Mapping

University of Iowa

451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu

The sequence contained an oligo-dT track that was present in the

oligonucleotide that was used to prime the synthesis of first

strand cDNA and therefore this may represent a bonafide poly A

tail. The sequence tag present in the cDNA between the NotI site

and the oligo-dT track served to identify it as a clone from the

normalized pons library cDNA Library Preparation: M.B. Soares Lab

Clone distribution: clones will be available through Research

Genetics (www.resgen.com)

Seq primer: M13 Forward

POLYA=Yes.

Location/Qualifiers

1. .568

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-CA1-bbs-e-06-0-UI"

/clone_lib="UI-R-CA1"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-CA1

library is a subtracted library derived from the following

tissues: thalamus, cerebellum, hypothalamus, medulla, pons

midbrain, cerebral cortex, corpus striatum, testis, and

hippocampus. For a detailed description of the library

from which this clone was derived, please visit our web

site at ratest.eng.uiowa.edu. The subtraction has been

previously described in (Bonaldo, Lennon and Soares,

Genome Research 6:791-806, 1996)

TAG LIB=UI-R-CA1

TAG_TISSUE=pons

TAG_SEQ=AGCAGC

BASE COUNT 116 a 156 c 131 g 165 t
ORIGIN

Query Match 18.2%; Score 64; DB 12; Length 568;
Best Local Similarity 87.5%; Pred. No. 3e-08;
Matches 70; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 272 CCRAAGGGAGGAGCGTCAGTTCAGTGGAAACAGGACGCTCGAGTGGGTGAGGTGT 331
|||||
Db 567 CAAGGGAGGAGGAGCGTCAGTTCAGTGGAAACAGGACGCTCGGTGGGTGAGGTGT 508
|||||

[illegible][illegible]

1

RESULT 14
BQ038901
LOCUS

GGAGCAGCACGGTCTCAATGGTCACAAACTTTCGCTC 626

B0038901 528 bp mRNA linear EST 01-MAY-2002

```
DEFINITION pgnlc.pk010.m22 normalized chicken lymphoid cdna library Gallus
gallus cDNA clone pgnlc.pk010.m22 5' similar to embjCAG84904.1
(AJ317960) transposition factor [Gallus gallus], mRNA sequence.
ACCESSION BQ038901
VERSION BQ038901.2 GI:20383663
KEYWORDS EST.
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
REFERENCE 1 (bases 1 to 528)
AUTHORS Morgan, R.W. and Burnside, J.
TITLE Chicken lymphoid ESTs
JOURNAL Unpublished (2001)
COMMENT On Mar 27, 2002 this sequence version replaced gi:19772441.
Contact: Joan Burnside
Molecular Endocrinology
University of Delaware
40 Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1345
Fax: 302-831-3411
Email: joan@udel.edu, www.chickest.udel.edu.
FEATURES
source
1..528
Location/Qualifiers
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="pgnlc.pk010.m22"
/clone.lib="normalized chicken lymphoid cdna library"
/sex="Male and Female"
/tissue_type="thymus, bursa, spleen, PBL, bone marrow"
/lab_host="E.coli EMDH10B"
/notes="Vector: pCMVSPORT 6"
BASE COUNT 110 a 202 c 124 g 79 t 13 others
ORIGIN
Query Match 9.9%; Score 34.6; DB 14; Length 528;
Best Local Similarity 44.2%; Pred. No. 15;
Matches 118; Conservative 0; Mismatches 149; Indels 0; Gaps 0;
QY 48 GCCGTACGGGCGCTCGAGAGCTGACACACATGCTGAACACACACACAGCATCTGTCT 107
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 244 GCTAGCGGCTCGGCTCCCGGACACAGCTCCACACCCACACCTCAGTGTGTCAT 303
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 108 CCAGTTTTCATTGGTGCAGGATCATGTCGATTTAGTTACTCTGACCCACAGCATCTGT 167
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 304 CAGCTCTGTCACATGCCACATCCACCGAGTTCGGGCTCTTTTTCACCAAGGCCAA 363
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 168 GTCATAGCCCAAGAACAAATACCGTGTATGATTCAGCTGGAGAAATAGAGCCCTTC 227
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 364 CAGCTATTCCACCACCACCACCATGATGATGATGATGATGATGATGATGATGATGAT 423
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 228 CAATGTGACGACATCATCCATCTCTGTACCTCCCGAGAGAGCCAAAGGGGAGAGCGT 287
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 424 CACAGTGGGACAACTCCAAAGTGCAGACGCCCCCAGNNNGCCCCCAGCGGACTCCCT 483
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 288 GCAGTTCAGTGGAAACAGGACAGCCT 314
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 484 GCAGAGCCAGGTCNNNNNGGNTGGCT 510
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
RESULT 15
BF135143/c
LOCUS 601780186F1 NCL_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4008333 5',
DEFINITION mRNA sequence.
ACCESSION BF135143
VERSION BF135143.1 GI:10974183
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

```
REFERENCE 1 (bases 1 to 917)
NTH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabbs@email.nih.gov
Tissue procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LIAW9242 row: k column: 22
High quality sequence stop: 708.
FEATURES
source
1..917
Location/Qualifiers
/organism="Mus musculus"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone.lib="IMAGE:4008333"
/clone.lib="NCL_CGAP_Lu30"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; transgenic model WNT-1, expression driven by
MTMV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
dT. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT 232 a 204 c 266 g 215 t
ORIGIN
Query Match 9.9%; Score 34.6; DB 12; Length 917;
Best Local Similarity 59.8%; Pred. No. 19;
Matches 58; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
QY 152 ACCCCAGCATCAGTGTGTATAGCCCAAGAACATACCGTGTGATTGATTGAGTGGAGA 211
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 188 ATCCAAGATTCTCTGGCTTGACGGGCAAGCACTTCACCAATGAACCTCAGTCATCCAGCA 129
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 212 AAATTAGAGCCCTTCCAAATGTGAGCACAGTCATCCA 248
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 128 TCCTTCCAGCCTCTTCCAGCAGTGAACCTCAGTCATCCA 92
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Search completed: November 6, 2002, 15:26:59
Job time : 2220 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 7, 2002, 04:21:31 ; Search time 801 seconds
(without alignments)
155.468 Million cell updates/sec

Title: US09897438BK-2

Perfect score: 351

Sequence: 1 GARCARTYGGNACNATNAT.....AYGARGCNTGTGGCNYTN 351

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 320260 seqs, 177392727 residues

Word size : 0

Total number of hits satisfying chosen parameters: 640520

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- Published_Applications_NA:*
- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
 - 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
 - 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
 - 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
 - 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
 - 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
 - 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
 - 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
 - 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
 - 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
 - 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
 - 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
 - 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
 - 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	4.6	2379	10	US-09-781-080B-12
2	15	4.3	696	10	US-09-810-052-3
3	15	4.3	729	10	US-09-810-052-6
4	14	4.0	792	10	US-09-390-017-4
5	14	4.0	1584	10	US-09-732-224-3
6	14	4.0	1629	10	US-09-732-224-6
7	14	4.0	1728	10	US-09-794-589-3
8	14	4.0	2094	10	US-09-899-471-6
9	13	3.7	51	10	US-09-853-253-18
10	13	3.7	693	10	US-09-728-911-3
11	13	3.7	696	10	US-09-825-561A-8
12	13	3.7	882	10	US-09-947-744-3
13	13	3.7	906	10	US-09-948-078-3
14	13	3.7	987	10	US-09-732-227-3
15	13	3.7	1644	10	US-09-819-136-7
16	13	3.7	1965	12	US-10-001-632A-3
17	13	3.7	2082	12	US-09-846-996A-2
18	12	3.4	26	10	US-09-986-676A-4
19	12	3.4	429	10	US-09-934-814-6

20	12	3.4	432	9	US-09-929-230-9	Sequence 9, Appli
21	12	3.4	435	10	US-09-923-995-3	Sequence 3, Appli
22	12	3.4	474	9	US-09-929-230-12	Sequence 12, Appli
23	12	3.4	519	10	US-09-745-003-5	Sequence 5, Appli
24	12	3.4	525	10	US-09-934-814-3	Sequence 3, Appli
25	12	3.4	537	10	US-09-745-003-3	Sequence 3, Appli
26	12	3.4	537	10	US-09-934-814-9	Sequence 9, Appli
27	12	3.4	597	10	US-09-801-231-3	Sequence 3, Appli
28	12	3.4	792	10	US-09-934-814-12	Sequence 12, Appli
29	12	3.4	975	10	US-09-741-711-3	Sequence 3, Appli
30	12	3.4	1038	10	US-09-122-383-13	Sequence 13, Appli
31	12	3.4	1059	9	US-09-911-345-3	Sequence 3, Appli
32	12	3.4	1140	12	US-10-003-356-6	Sequence 6, Appli
33	12	3.4	1200	10	US-09-921-823-3	Sequence 3, Appli
34	12	3.4	1245	12	US-10-005-947-3	Sequence 3, Appli
35	12	3.4	1833	10	US-09-921-823-18	Sequence 18, Appli
36	12	3.4	1833	10	US-09-921-823-21	Sequence 21, Appli
37	12	3.4	2022	10	US-09-899-471-3	Sequence 3, Appli
38	12	3.4	2337	10	US-09-731-179-3	Sequence 9, Appli
39	12	3.4	2781	12	US-10-003-356-9	Sequence 3, Appli
40	12	3.4	4527	10	US-09-901-940-3	Sequence 3, Appli
41	11	3.1	24	10	US-09-865-018-13	Sequence 13, Appli
42	11	3.1	252	10	US-09-733-523-3	Sequence 3, Appli
43	11	3.1	333	10	US-09-750-964-3	Sequence 3, Appli
44	11	3.1	426	10	US-09-781-077-6	Sequence 6, Appli
45	11	3.1	435	10	US-09-740-638-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-781-080B-12
; Sequence 12, Application US/09781080B
; Patent No. US20020142439A1
; GENERAL INFORMATION:
; APPLICANT: HOLLOWAY, JIM
; APPLICANT: SHEPPARD, PAUL
; APPLICANT: YAMAMOTO, GAYLE
; TITLE OF INVENTION: Anti-Angiogenic Intestinal Peptides,
; FILE REFERENCE: 99-82
; CURRENT APPLICATION NUMBER: US/09/781.080B
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: fastseq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 2379
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Degenerate sequence
; NAME/KEY: misc_feature
; LOCATION: (1)...(2379)
; OTHER INFORMATION: n = A,T,C or G
US-09-781-080B-12

Query Match 4.6%; Score 16; DB 10; Length 2379;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 NWSNGNNSNTGYMGN 138

Db 1512 NWSNGNNSNTGYMGN 1527

RESULT 2

US-09-810-052-3
; Sequence 3, Application US/09810052
; Patent No. US20020009775A1
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Presnell, Scott R.

```
; TITLE OF INVENTION: HELICAL PROTEIN ZALPHA51
; FILE REFERENCE: 00-24
; CURRENT APPLICATION NUMBER: US/09/810,052
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/790,410
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/199,443
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 696
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: degenerate sequence
; NAME/KEY: misc_feature
; LOCATION: (1)...(696)
; OTHER INFORMATION: n = A,T,C or G
US-09-810-052-3

Query Match
Best Local Similarity 4.3%; Score 15; DB 10; Length 696;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 AYTNCNGARGARG 271
|||||
Db 449 AYTNCNGARGARG 463

RESULT 3
US-09-810-052-6
; Sequence 6, Application US/09810052
; Patent No. US20020009775A1
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; TITLE OF INVENTION: HELICAL PROTEIN ZALPHA51
; FILE REFERENCE: 00-24
; CURRENT APPLICATION NUMBER: US/09/810,052
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/199,443
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/199,443
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 729
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: degenerate sequence
; NAME/KEY: misc_feature
; LOCATION: (1)...(729)
; OTHER INFORMATION: n = A,T,C or G
US-09-810-052-6

Query Match
Best Local Similarity 4.3%; Score 15; DB 10; Length 729;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 AYTNCNGARGARG 271
|||||
Db 482 AYTNCNGARGARG 496

RESULT 4
US-09-990-017-4
; Sequence 4, Application US/09990017
; Patent No. US20020115168A1
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
```

```
; APPLICANT: Gao, Zeren
; TITLE OF INVENTION: NOVEL PROTEIN ZLMDA2
; FILE REFERENCE: 00-67
; CURRENT APPLICATION NUMBER: US/09/990,017
; CURRENT FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: US 60/252,374
; PRIOR FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 792
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: degenerate sequence
; NAME/KEY: misc_feature
; LOCATION: 12, 15, 36, 39, 42, 45, 54, 60, 63, 66, 75, 78, 84, 93, 99,
; LOCATION: 102, 105, 108, 111, 114, 117, 120, 123, 126, 144, 147, 153,
; LOCATION: 162, 168, 174, 180, 183, 186, 189, 198, 204, 210, 213, 219,
; LOCATION: 222, 228, 231, 234, 240, 246, 249, 252, 258, 261, 264
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: 273, 279, 282, 294, 300, 303, 306, 309, 321, 327, 345, 372,
; LOCATION: 378, 384, 393, 396, 399, 402, 405, 408, 423, 426, 438, 444,
; LOCATION: 447, 450, 468, 474, 483, 486, 498, 504, 510, 513, 516, 525,
; LOCATION: 528, 531, 534, 537, 543, 549, 552, 564, 585, 591, 594
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: 597, 618, 630, 648, 651, 654, 657, 660, 663, 678, 696, 699,
; LOCATION: 702, 705, 708, 711, 714, 717, 720, 723, 729, 738, 750, 756,
; LOCATION: 765, 771, 777, 780, 783, 789, 792
; OTHER INFORMATION: n = A,T,C or G
US-09-990-017-4

Query Match
Best Local Similarity 4.0%; Score 14; DB 10; Length 792;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 215 TMNGNGCNCNWSN 228
|||||
Db 647 TMNGNGCNCNWSN 660

RESULT 5
US-09-732-224-3
; Sequence 3, Application US/09732224
; Patent No. US20020095022A1
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: Human Secretin-Like Receptor
; FILE REFERENCE: 99-100
; CURRENT APPLICATION NUMBER: US/09/732,224
; CURRENT FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1584
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: This degenerate nucleotide sequence encodes the
; OTHER INFORMATION: amino acid sequence of SEQ ID NO:2.
; NAME/KEY: variation
; LOCATION: (1)...(1584)
; OTHER INFORMATION: N is any nucleotide.
US-09-732-224-3

Query Match
Best Local Similarity 4.0%; Score 14; DB 10; Length 1584;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 GNCCNMGNGARYTN 69
|||||
```

Db 374 GNCNMGNARYTN 387

RESULT 6

; Sequence 6, Application US/09732224
; Patent No. US20020095022A1

; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.

; TITLE OF INVENTION: Human Secretin-Like Receptor
; FILE REFERENCE: 99-100

; CURRENT APPLICATION NUMBER: US/09/732,224
; CURRENT FILING DATE: 2000-12-07

; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 6
; LENGTH: 1629

; TYPE: DNA
; ORGANISM: Artificial Sequence

; FEATURE:
; OTHER INFORMATION: This degenerate nucleotide sequence encodes the

; NAME/KEY: variation
; LOCATION: (1)...(1629)

; OTHER INFORMATION: N is any nucleotide.
US-09-732-224-6

Query Match 4.0%; Score 14; DB 10; Length 1629;

Best Local Similarity 100.0%; Pred. No. 0.0018;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 56 GNCNMGNARYTN 69

Db 374 GNCNMGNARYTN 387

RESULT 7

US-09-794-589-3

; Sequence 3, Application US/09794589

; Patent No. US20020004224A1

; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.

; TITLE OF INVENTION: KUNITZ DOMAIN POLYPEPTIDE ZKUN8
; FILE REFERENCE: 00-01

; CURRENT APPLICATION NUMBER: US/09/794,589
; CURRENT FILING DATE: 2001-02-27

; PRIOR APPLICATION NUMBER: US 60/186,069
; PRIOR FILING DATE: 2000-02-29

; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 3
; LENGTH: 1728

; TYPE: DNA
; ORGANISM: Artificial Sequence

; FEATURE:
; OTHER INFORMATION: degenerate sequence

; NAME/KEY: misc_feature
; LOCATION: (1)...(1728)

; OTHER INFORMATION: n = A,T,C or G
US-09-794-589-3

Query Match 4.0%; Score 14; DB 10; Length 1728;

Best Local Similarity 100.0%; Pred. No. 0.0018;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 CNACNACNGCNSN 102

Db 584 CNACNACNGCNSN 597

RESULT 8

US-09-899-471-6

; Sequence 6, Application US/09899471

; Patent No. US20020146763A1
; GENERAL INFORMATION:
; APPLICANT: Gao, Zeren

; TITLE OF INVENTION: Murine Cytokine Receptor
; FILE REFERENCE: 00-46

; CURRENT APPLICATION NUMBER: US/09/899,471
; CURRENT FILING DATE: 2001-07-05

; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 6
; LENGTH: 2094

; TYPE: DNA
; ORGANISM: Artificial Sequence

; FEATURE:
; OTHER INFORMATION: This degenerate nucleotide sequence encodes the

; NAME/KEY: misc_feature
; LOCATION: (1)...(2094)

; OTHER INFORMATION: n = A,T,C or G
US-09-899-471-6

Query Match 4.0%; Score 14; DB 10; Length 2094;

Best Local Similarity 100.0%; Pred. No. 0.0018;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 228 NAAyGTNWSNACNG 241

Db 624 NAAyGTNWSNACNG 637

RESULT 9

US-09-853-253-18

; Sequence 18, Application US/09853253

; Patent No. US20020055156A1

; GENERAL INFORMATION:
; APPLICANT: JASPER, STEPHEN

; APPLICANT: SHEPPARD, PAUL
; APPLICANT: DEISHER, THERESA

; APPLICANT: BISHOP, PAUL
; TITLE OF INVENTION: zsig33-like Peptides

; FILE REFERENCE: 00-30
; CURRENT APPLICATION NUMBER: US/09/853,253

; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: 60/203,300

; PRIOR FILING DATE: 2000-05-11
; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18

; LENGTH: 51
; TYPE: DNA

; ORGANISM: Artificial Sequence
; FEATURE:

; OTHER INFORMATION: degenerate sequence
; NAME/KEY: misc_feature

; LOCATION: (1)...(51)
; OTHER INFORMATION: n = A,T,C or G

US-09-853-253-18

Query Match 3.7%; Score 13; DB 10; Length 51;

Best Local Similarity 100.0%; Pred. No. 0.026;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 265 GARGCNCNARG 277

Db 22 GARGCNCNARG 34

RESULT 10

US-09-728-911-3

; Sequence 3, Application US/09728911

; Patent No. US20020012669A1

; GENERAL INFORMATION:
; APPLICANT: Presnell, Scott R.

APPLICANT: Xu, Wenfeng
APPLICANT: Kindsvogel, Wayne
APPLICANT: Chen, Zhi
TITLE OF INVENTION: Human Cytokine Receptor
FILE REFERENCE: 99-93
CURRENT APPLICATION NUMBER: US/09/728,911
CURRENT FILING DATE: 2000-12-01
PRIOR APPLICATION NUMBER: US 60/169,049
PRIOR FILING DATE: 1999-12-03
PRIOR APPLICATION NUMBER: US 60/232,219
PRIOR FILING DATE: 2000-09-13
PRIOR APPLICATION NUMBER: US 60/244,610
PRIOR FILING DATE: 2000-10-31
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 3
LENGTH: 693
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: This degenerate nucleotide sequence encodes the
OTHER INFORMATION: amino acid sequence of SEQ ID NO:2.
NAME/KEY: misc_feature
LOCATION: (1)...(693)
OTHER INFORMATION: n = A,T,C or G
US-09-728-911-3

Query Match 3.7%: Score 13; DB 10; Length 693;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 285 NGTNCARTTCAR 297
|||||
Db 99 NGTNCARTTCAR 111

RESULT 11

US-09-825-561A-8
Sequence 8, Application US/09825561A
Patent No. US2002013767A1
GENERAL INFORMATION:
APPLICANT: Sprecher, Cindy A.
APPLICANT: West, James W.
APPLICANT: Presnell, Scott R.
APPLICANT: Holly, Richard D.
APPLICANT: Nelson, Andrew J.
TITLE OF INVENTION: SOLUBLE ZALPHAL1 CYTOKINE RECEPTORS
FILE REFERENCE: 00-22
CURRENT APPLICATION NUMBER: US/09/825,561A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/194,731
PRIOR FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/222,121
PRIOR FILING DATE: 2000-07-28
NUMBER OF SEQ ID NOS: 86
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 8
LENGTH: 696
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Degenerate polynucleotide sequence of IL-2Rgamma
OTHER INFORMATION: polypeptide as shown in SEQ ID NO:4
NAME/KEY: misc_feature
LOCATION: (1)...(696)
OTHER INFORMATION: n = A,T,C or G
US-09-825-561A-8

Query Match 3.7%: Score 13; DB 10; Length 696;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 82 YTNAAYACNACNA 94
|||||
Db 1 YTNAAYACNACNA 13

RESULT 12

US-09-947-744-3
Sequence 3, Application US/09947744
Patent No. US20020146766A1
GENERAL INFORMATION:
APPLICANT: Holloway, James L.
APPLICANT: Lok, Si
TITLE OF INVENTION: Human Vomeronal Receptor-3
FILE REFERENCE: 00-82
CURRENT APPLICATION NUMBER: US/09/947,744
CURRENT FILING DATE: 2001-09-06
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 3
LENGTH: 882
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: This degenerate nucleotide sequence encodes the
OTHER INFORMATION: amino acid sequence of SEQ ID NO:2.
NAME/KEY: misc_feature
LOCATION: 21, 24, 27, 36, 42, 48, 51, 57, 63, 66, 72, 78, 81, 87, 93,
LOCATION: 96, 99, 111, 114, 120, 123, 132, 135, 141, 144, 150, 156,
LOCATION: 165, 168, 171, 186, 189, 192, 195, 198, 210, 213, 225, 228,
LOCATION: 237, 261, 264, 273, 279, 282, 288, 291, 294, 297, 303
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc_feature
LOCATION: 306, 309, 315, 318, 321, 327, 330, 333, 339, 345, 348, 351,
LOCATION: 354, 357, 363, 366, 369, 378, 384, 396, 399, 408, 414, 420,
LOCATION: 432, 438, 444, 450, 453, 456, 459, 462, 465, 477, 480, 483,
LOCATION: 489, 492, 495, 501, 504, 510, 516, 519, 522, 528, 531
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc_feature
LOCATION: 540, 549, 558, 561, 567, 579, 582, 585, 588, 591, 594, 597,
LOCATION: 600, 603, 606, 609, 612, 621, 627, 630, 633, 636, 639, 645,
LOCATION: 651, 666, 669, 675, 678, 681, 687, 690, 693, 705, 708, 714,
LOCATION: 720, 723, 726, 729, 732, 741, 747, 753, 756, 765, 768
OTHER INFORMATION: n = A,T,C or G
NAME/KEY: misc_feature
LOCATION: 771, 774, 780, 783, 786, 789, 798, 804, 807, 810, 816, 819,
LOCATION: 828, 831, 840, 846, 849, 855, 858, 861, 864, 867, 870, 873,
LOCATION: 876
OTHER INFORMATION: n = A,T,C or G
US-09-947-744-3

Query Match 3.7%: Score 13; DB 10; Length 882;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 72 NACNACNTGYTN 84
|||||
Db 303 NACNACNTGYTN 315

RESULT 13

US-09-948-078-3
Sequence 3, Application US/09948078
Patent No. US20020147308A1
GENERAL INFORMATION:
APPLICANT: Holloway, James L.
APPLICANT: Lok, Si
TITLE OF INVENTION: Human Vomeronal Receptor
FILE REFERENCE: 00-73
CURRENT APPLICATION NUMBER: US/09/948,078
CURRENT FILING DATE: 2001-09-06
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 3


```

; LENGTH: 906
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: This degenerate nucleotide sequence encodes the
; OTHER INFORMATION: amino acid sequence of SEQ ID NO:2.
; NAME/KEY: misc_feature
; LOCATION: 6, 9, 12, 18, 21, 24, 27, 39, 42, 48, 51, 54, 57, 60, 63,
; LOCATION: 66, 69, 75, 78, 81, 84, 93, 96, 99, 108, 111, 117, 120,
; LOCATION: 123, 126, 129, 135, 141, 150, 156, 159, 168, 171, 174, 177,
; LOCATION: 186, 189, 192, 198, 204, 207, 213, 216, 219, 228, 234, 237
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: 240, 249, 252, 261, 267, 270, 273, 276, 279, 282, 285, 291,
; LOCATION: 294, 297, 303, 306, 309, 312, 321, 327, 330, 333, 336, 339,
; LOCATION: 345, 348, 354, 360, 372, 375, 384, 387, 393, 396, 399, 402,
; LOCATION: 414, 423, 426, 441, 450, 453, 456, 471, 480, 483, 495
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: 498, 507, 510, 513, 516, 531, 537, 540, 543, 546, 552, 555,
; LOCATION: 558, 564, 570, 573, 579, 582, 585, 591, 597, 600, 603, 606,
; LOCATION: 612, 621, 627, 639, 642, 657, 660, 666, 669, 672, 675, 678,
; LOCATION: 681, 684, 693, 696, 699, 705, 711, 717, 720, 723, 726
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc_feature
; LOCATION: 732, 735, 738, 744, 747, 750, 756, 765, 768, 774, 777, 780,
; LOCATION: 789, 795, 798, 801, 804, 810, 813, 816, 819, 825, 828, 837,
; LOCATION: 840, 843, 846, 849, 855, 858, 864, 873, 876, 879, 885, 888,
; LOCATION: 897, 906
; OTHER INFORMATION: n = A,T,C or G
US-09-948-078-3
```

```

Query Match          3.7%; Score 13; DB 10; Length 906;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 72 NACNACNTGYTGN 84
Db 291 NACNACNTGYTGN 303
```

```

RESULT 14
US-09-732-227-3
; Sequence 3, Application US/09732227
; Patent No. US20020004228A1
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: zwf1: A Member of the Von Willebrand
; TITLE OF INVENTION: Factor Type A Domain Superfamily
; FILE REFERENCE: 99-99
; CURRENT APPLICATION NUMBER: US/09/732.227
; CURRENT FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 987
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: This degenerate nucleotide sequence encodes the
; OTHER INFORMATION: amino acid sequence of SEQ ID NO:2.
; NAME/KEY: variation
; LOCATION: (1)...(987)
; OTHER INFORMATION: N is any nucleotide.
US-09-732-227-3
```

```

Query Match          3.7%; Score 13; DB 10; Length 987;
Best Local Similarity 100.0%; Pred. No. 0.016;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 102 NGTNYTNCARTTY 114
Db 675 NGTNYTNCARTTY 687
```

```

RESULT 15
US-09-819-136-7
; Sequence 7, Application US/09819136
; Patent No. US20020146789A1
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Gao, Zeren
; TITLE OF INVENTION: MULTI-DOMAIN PROTEINASE INHIBITOR
; FILE REFERENCE: 00-25
; CURRENT APPLICATION NUMBER: US/09/819.136
; CURRENT FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: US 60/193,642
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 1644
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: degenerate sequence
; NAME/KEY: misc_feature
; LOCATION: (1)...(1644)
; OTHER INFORMATION: n = A,T,C or G
US-09-819-136-7
```

```

Query Match          3.7%; Score 13; DB 10; Length 1644;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 330 NTAYGARGCNTGY 342
Db 1029 NTAYGARGCNTGY 1041
```

```

Search completed: November 7, 2002, 06:27:12
Job time : 805 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 6, 2002, 23:00:45 ; Search time 89 Seconds
(without alignments)
437.968 Million cell updates/sec

Title: US-09-897-438B-2

Perfect score: 117
Sequence: 1 EQCGTTHGNNAVTFCEPYGP.....FQWKQSLRGEVVEACWAL 117

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 310279 seqs, 166577418 residues

Word size: 1

Total number of hits satisfying chosen parameters: 619741

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlp
-Q/cgn2_1/USPTO.spool/US09897438/runat_06112002_101958/app_query.fasta_1.263
-DB=Published Applications_NA -QFMT=fastap -SURFIX=oligo.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-TRANS=human40.cdi -LISN=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09897438 -ECGN_1_1_36=runat_06112002_101958_11894
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Published Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
117	100.0	351	10	US-09-897-438B-1	Sequence 1, Appli
27	23.1	11580	10	US-09-880-107-3436	Sequence 3436, Ap
18	15.4	413	10	US-09-864-761-11808	Sequence 11808, A
14	12.0	87	10	US-09-864-761-28379	Sequence 28379, A

c 5	13	11.1	214	10	US-09-864-761-31037	Sequence 31037, A
c 6	13	11.1	499	10	US-09-864-761-14497	Sequence 14497, A
c 7	8	6.8	1132	12	US-10-014-927-20	Sequence 20, Appl
c 8	7	6.0	21	10	US-09-897-438B-10	Sequence 10, Appl
c 9	7	6.0	22	10	US-09-897-438B-6	Sequence 6, Appl
c 10	7	6.0	31	10	US-09-894-698-14	Sequence 14, Appl
c 11	7	6.0	238	10	US-09-783-590-2199	Sequence 2199, Ap
c 12	7	6.0	239	10	US-09-960-352-11910	Sequence 11910, A
c 13	7	6.0	365	10	US-09-864-761-15576	Sequence 15576, A
c 14	7	6.0	404	10	US-09-833-381-647	Sequence 647, App
c 15	7	6.0	637	10	US-09-922-217-945	Sequence 945, App
c 16	7	6.0	637	10	US-09-833-263-945	Sequence 945, App
c 17	7	6.0	669	10	US-09-864-761-18014	Sequence 18014, A
c 18	7	6.0	880	10	US-09-770-445-531	Sequence 531, App
c 19	7	6.0	896	10	US-09-908-805B-28	Sequence 28, Appl
c 20	7	6.0	1008	10	US-09-815-242-8010	Sequence 8010, Ap
c 21	7	6.0	1221	10	US-09-974-300-4246	Sequence 4246, Ap
c 22	7	6.0	2000	10	US-09-887-576-870	Sequence 870, App
c 23	7	6.0	2095	10	US-09-764-877-3948	Sequence 3948, Ap
c 24	7	6.0	2343	10	US-09-810-264-23	Sequence 23, Appl
c 25	7	6.0	2346	10	US-09-815-242-7180	Sequence 7180, Ap
c 26	7	6.0	2608	10	US-09-739-254-16	Sequence 16, Appl
c 27	7	6.0	2608	10	US-09-904-615-16	Sequence 16, Appl
c 28	7	6.0	2628	10	US-09-815-242-4377	Sequence 4377, Ap
c 29	7	6.0	2634	10	US-09-815-242-8458	Sequence 8458, Ap
c 30	7	6.0	2711	10	US-09-764-877-3950	Sequence 3950, Ap
c 31	7	6.0	2712	10	US-09-764-877-3949	Sequence 3949, Ap
c 32	7	6.0	3719	10	US-09-917-800A-1705	Sequence 1705, Ap
c 33	7	6.0	13968	10	US-09-764-869-2224	Sequence 2224, Ap
c 34	7	6.0	16854	10	US-09-764-878-217	Sequence 217, App
c 35	7	6.0	19736	12	US-10-014-502-3	Sequence 3, Appl
c 36	7	6.0	98865	10	US-09-770-689A-3	Sequence 3, Appl
c 37	6	5.1	18	10	US-09-854-799-39	Sequence 39, Appl
c 38	6	5.1	18	10	US-09-897-438B-11	Sequence 11, Appl
c 39	6	5.1	80	10	US-09-864-761-21132	Sequence 21132, A
c 40	6	5.1	92	10	US-09-793-306-156	Sequence 156, App
c 41	6	5.1	95	10	US-09-864-761-30980	Sequence 30980, A
c 42	6	5.1	95	10	US-09-878-574-6384	Sequence 6384, Ap
c 43	6	5.1	107	10	US-09-878-574-6384	Sequence 6384, Ap
c 44	6	5.1	112	10	US-09-864-761-20139	Sequence 20139, A
c 45	6	5.1	119	10	US-09-294-093B-2470	Sequence 2470, Ap

ALIGNMENTS

RESULT 1
US-09-897-438B-1
; Sequence 1, Application US/09897438B
; Patent No. US20020137095A1
; GENERAL INFORMATION:
; APPLICANT: Tate, Naoko
; APPLICANT: Mikoshiba, Katsuhiko
; TITLE OF INVENTION: REELIN PROTEIN CR-50 EPTOPE REGION
; FILE REFERENCE: 04853-0076-00000
; CURRENT APPLICATION NUMBER: US/09/897,438B
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: JP 2000-202801
; PRIOR FILING DATE: 2000-07-04
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 351
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-897-438B-1

Self

Alignment Scores:
Pred. No.: 8.82e-121 Length: 351
Score: 117.00 Matches: 117
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

```

US-09-897-438B-2 (1-117) x US-09-897-438B-1 (1-351)
Qy 1 GluGlnCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20
Db 1 GAGCAGTGTGGCACCACATCATGCGCAATGCTGTCACCTTCTGTGAGCCGTACGGCCCT 60
Qy 21 ArgGluLeuThrThrThrCysLeuAsnThrThrThrAlaSerValLeuGlnPheSerIle 40
Db 61 CGAGAGCTGACCACACATGCTGAACACAAACAGCATCTGCTCCAGTTTCCATT 120
Qy 41 GlySerGlySerCysArgPheSerTyrSerAspProSerIleThrValSerTyrAlaLys 60
Db 121 GGGTCAGGATCATGTCGATTAGTTACTCTGACCCCAACATCACTGTGTATGACGCCAAG 180
Qy 61 AsnAsnThrAlaAspTrpIleGlnLeuGluLysIleArgAlaProSerAsnValSerThr 80
Db 181 AACATACCGCTGATTGATTCAGTTCAGTGGAGAAATAGAGCCCTTCCATGTGAGCACA 240
Qy 81 ValIleHisIleLeuTyrLeuProGluGluAlaLysGlyGluSerValGlnPheGlnTrp 100
Db 241 GTCATCCACATCTGTACCTCCCGAGGAGCAAGAGGGAGAGCGTGCAGTTCCAGTGG 300
Qy 101 LysGlnAspSerLeuArgValGlyGluValTyrGluAlaCysTrpAlaLeu 117
Db 301 AAACAGGACAGCTGCGAGTGGGTGAGGTGTATGAGGCTGCTGGGCCCTG 351

```

RESULT 2

```

US-09-880-107-3436
; Sequence 3436, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; EMBL REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3436
; LENGTH: 11580
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U79716
US-09-880-107-3436

```

```

Alignment Scores:
Pred. No.: 6,48e-20 Length: 11580
Score: 27.00 Matches: 27
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 23.08% Indels: 0
DB: 10 Gaps: 0

```

US-09-897-438B-2 (1-117) x US-09-880-107-3436 (1-11580)

```

Qy 28 LeuAsnThrThrThrAlaSerValLeuGlnPheSerIleGlySerGlySerCysArgPhe 47
Db 941 CTTAATACACACAGCTTCTGCTCCATTTTCCATTGGGTGAGCGTTCATGTCGCTTT 1000
Qy 48 SerTyrSerAspProSerIle 54
Db 1001 AGTTATTACAGCCCGCCAGCATC 1021

```

RESULT 3

```

US-09-864-761-11808/c
; Sequence 11808, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Hanzel, David R.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Acomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anncmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 11808
; LENGTH: 413
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC000121.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.55
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.57
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.51
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.44
US-09-864-761-11808

```

```

Alignment Scores:
Pred. No.: 2,14e-11 Length: 413
Score: 18.00 Matches: 18
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 15.38% Indels: 0
DB: 10 Gaps: 0

```

US-09-897-438B-2 (1-117) x US-09-864-761-11808 (1-413)

US-09-864-761-31037

Alignment Scores:		
Pred. No.:	3.72e-06	214
Score:	13.00	13
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	11.11%	Indels: 0
DB:	10	Gaps: 0

US-09-897-438B-2 (1-117) x US-09-864-761-31037 (1-214)

Qy 105 LeuArgValGlyGluValTyrGluAlaCysTrpAlaLeu 117
 |||||
 Db 137 CTTCTGTGAGGTGAAGTGTATGAAGCCTGCTGGGCCTTA 99

```

; OTHER INFORMATION: MAP TO AC000121.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.62
US-09-864-761-14497

```

Alignment Scores:	8.53e-06	Length:	499
Pred. No.:	13.00	Matches:	13
Score:	100.00%	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	11.11%	Gaps:	0
DB:	10		

US-09-897-438B-2 (1-117) X US-09-864-761-14497 (1-499)

105 LeuArgValGlvGluValTyrGluAlaCysTrpAlaLeu 117

```

; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence:atSRp34/SR1
MS-10-014-927-20

```

Alignment Scores:		
Pred. No.:	6.32	Length: 1132
Score:	8.00	Matches: 8
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	6.84%	Indels: 0
DB:	12	Gaps: 0

US-09-897-438B-2 (1-117) X US-10-014-927-20 (1-1132)

Qy 76 SerAsnValSerThrValIleHis 83
 |||||
Db 593 TCCAATGTCTCCACTGTTATCCAT 570

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: primer for PCR
US-09-897-438B-10

Alignment Scores:
Pred. No.: 1.61 Length: 21
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.98% Indels: 0
DB: 1.0 Gaps: 0

US-09-897-438B-2 (1-117) x US-09-897-438B-10 (1-21)

Qy 1 GluInCysGlyThrIleMet 7
Db 1 GAGCAGTGTGGCACCATCATG 21

RESULT 9
US-09-897-438B-6/C
; Sequence 6, Application US/09897438B
; Patent No. US20020137095A1
; GENERAL INFORMATION:
; APPLICANT: Mikoshiba, Katsuhiko
; APPLICANT: Tate, Naoko
; TITLE OF INVENTION: REELIN PROTEIN CR-50 EPIPOPE REGION
; FILE REFERENCE: 04853-0076-00000
; CURRENT APPLICATION NUMBER: US/09/897.438B
; CURRENT FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: JP 2000-202801
; PRIOR FILING DATE: 2000-07-04
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: primer for PCR
US-09-897-438B-6

Alignment Scores:
Pred. No.: 1.69 Length: 22
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.98% Indels: 0
DB: 1.0 Gaps: 0

US-09-897-438B-2 (1-117) x US-09-897-438B-6 (1-22)

Qy 111 TyrGluAlaCysTrpAlaLeu 117
Db 21 TATGAGGCGCTGCTGGCCCTG 1

RESULT 10
US-09-894-698-14/C
; Sequence 14, Application US/09894698
; Patent No. US20020026041A1
; GENERAL INFORMATION:
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Wisniewski, Nancy
; TITLE OF INVENTION: FLEA ALLANTOINASE NUCLEIC ACID MOLECULES, PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: FC-6-cl-cl
; CURRENT APPLICATION NUMBER: US/09/894.698
```

```
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 09/543,668
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,704
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 31
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Primer
US-09-894-698-14

Alignment Scores:
Pred. No.: 2.36 Length: 31
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.98% Indels: 0
DB: 1.0 Gaps: 0

US-09-897-438B-2 (1-117) x US-09-894-698-14 (1-31)

Qy 36 LeuGlnPheSerIleGlySer 42
Db 23 TTGCAATTGAGCATAGGATCC 3

RESULT 11
US-09-783-590-2199/c
; Sequence 2199, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783.590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2199
; LENGTH: 238
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (22)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (43)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (93)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (117)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (121)
; OTHER INFORMATION: n equals a,t,g, or c
```

```

; NAME/KEY: misc feature
; LOCATION: (183)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (198)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (207)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (210)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (213)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (223)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (224)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-2199

Alignment Scores:
Pred. No.:      17.4      Length:      238
Score:          7.00     Matches:       7
Percent Similarity: 100.00% Conservative:    0
Best Local Similarity: 100.00% Mismatches:     0
Query Match:      5.98% Indels:         0
DB:              10      Gaps:         0

US-09-897-438B-2 (1-117) x US-09-783-590-2199 (1-238)

QY   17 ProTyrglyProArgGluLeu 23
Db   154 CCTATGGTCCCAGAGACTT 134

RESULT 12
US-09-960-352-11910/c
; Sequence 11910, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 11910
; LENGTH: 239
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 51-LIB3057-021-Q1-K1-E4
US-09-960-352-11910

Alignment Scores:
Pred. No.:      17.5      Length:      239
Score:          7.00     Matches:       7
Percent Similarity: 100.00% Conservative:    0
Best Local Similarity: 100.00% Mismatches:     0
Query Match:      5.98% Indels:         0
DB:              10      Gaps:         0

US-09-897-438B-2 (1-117) x US-09-960-352-11910 (1-239)

QY   44 SerCysargPheSerTyrSer 50
Db   153 TCATGTAGATTTCAGTTACACT 133

```



```
Db 44 ATTCAGTTAGAAAAATACGT 64
|||||
RESULT 14
US-09-833-381-647/c
; Sequence 647, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: NO. US20020132090A1el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 647
; LENGTH: 404
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(404)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-647

Alignment Scores:
Pred. No.: 29.3 Length: 404
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.98% Indels: 0
DB: 10 Gaps: 0

US-09-897-438B-2 (1-117) x US-09-833-381-647 (1-404)
Qy 42 SerGlySerCysArgPheSer 48
|||||
Db 202 TCGGGAAGCTGCAGATTTCAGT 182
|||||
RESULT 15
US-09-922-217-945
; Sequence 945, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922,217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 945
; LENGTH: 637
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 629
; OTHER INFORMATION: n = A,T,C or G

US-09-922-217-945
Alignment Scores:
Pred. No.: 45.7 Length: 637
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.98% Indels: 0
DB: 10 Gaps: 0

US-09-897-438B-2 (1-117) x US-09-922-217-945 (1-637)
Qy 42 SerGlySerCysArgPheSer 48
|||||
Db 372 TCTGGCAGTTGCCCTTCAGT 392
|||||

Search completed: November 7, 2002, 00:48:12
Job time : 91 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: November 6, 2002, 22:05:15 ; Search time 68 seconds
(without alignments)
527.664 Million cell updates/sec

Title: US-09-897-438B-2
Perfect score: 117
Sequence: 1 EOCGTIMHGNVATFCEPYGP.....FONKQDSLVRGVEVACWAL 117

Scoring table: OLIGO
Xgapop 60.0, Xgapext 60.0
Ygapop 60.0, Ygapext 60.0
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 153338381 residues

Word size: 1

Total number of hits satisfying chosen parameters: 878600

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool/US09897438/runat_06112002_101956_11823/app_query.fasta_1.263
-DB=Issued_Patents_NA -QPMT=fastap -SUFFIX=oligo.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HPAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09897438 -CGN_1_1_31_runat_06112002_101956_11823 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAPP -LARGQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_NA: *
1: /cgn2.6/ptodata/1/ina/5A.COMB.seq: *
2: /cgn2.6/ptodata/1/ina/5B.COMB.seq: *
3: /cgn2.6/ptodata/1/ina/6A.COMB.seq: *
4: /cgn2.6/ptodata/1/ina/6B.COMB.seq: *
5: /cgn2.6/ptodata/1/ina/PCTUS.COMB.seq: *
6: /cgn2.6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB ID	Description
1	117	100.0	11673	4	US-09-334-220-3	Sequence 3, Appli
2	27	23.1	11580	4	US-09-334-220-4	Sequence 4, Appli
3	7	6.0	31	4	US-09-894-698-14	Sequence 14, Appli
4	7	6.0	896	4	US-09-302-769-28	Sequence 28, Appli
5	7	6.0	1125	4	US-09-134-001C-2718	Sequence 2718, Ap
6	7	6.0	1200	4	US-09-222-938A-6	Sequence 6, Appli
7	7	6.0	1276	4	US-08-858-207A-109	Sequence 109, App
8	7	6.0	1666	4	US-09-221-017B-933	Sequence 933, App
9	7	6.0	1851	4	US-09-608-790-2	Sequence 2, Appli
10	7	6.0	1923	1	US-08-077-939-14	Sequence 14, Appli
11	7	6.0	1923	1	US-08-461-599-14	Sequence 14, Appli
12	7	6.0	1923	1	US-08-461-621-14	Sequence 14, Appli

c 13	7	6.0	1923	1	US-08-465-334-14	Sequence 14, Appli
c 14	7	6.0	1974	1	US-08-077-939-16	Sequence 16, Appli
c 15	7	6.0	1974	1	US-08-461-599-16	Sequence 16, Appli
c 16	7	6.0	1974	1	US-08-461-621-16	Sequence 16, Appli
c 17	7	6.0	1974	1	US-08-465-334-16	Sequence 16, Appli
c 18	7	6.0	2631	1	US-08-785-071A-1	Sequence 1, Appli
c 19	7	6.0	2631	3	US-09-012-872-1	Sequence 3, Appli
c 20	7	6.0	3057	1	US-08-551-459-3	Sequence 14, Appli
c 21	7	6.0	3250	4	US-09-122-126B-14	Sequence 5, Appli
c 22	7	6.0	3466	1	US-08-551-459-5	Sequence 7, Appli
c 23	7	6.0	4526	4	US-09-424-283-7	Sequence 3, Appli
c 24	7	6.0	19736	4	US-09-740-035-3	Sequence 39, Appli
c 25	6	5.1	18	4	US-08-392-459-39	Sequence 39, Appli
c 26	6	5.1	18	5	PCT-US91-08525-39	Sequence 20, Appli
c 27	6	5.1	25	1	US-07-884-811-20	Sequence 20, Appli
c 28	6	5.1	25	1	US-07-885-971-20	Sequence 20, Appli
c 29	6	5.1	25	1	US-08-087-783A-20	Sequence 20, Appli
c 30	6	5.1	25	1	US-08-194-088B-20	Sequence 20, Appli
c 31	6	5.1	25	1	US-08-435-501-19	Sequence 19, Appli
c 32	6	5.1	25	1	US-08-435-764-19	Sequence 19, Appli
c 33	6	5.1	25	1	US-08-792-078-19	Sequence 19, Appli
c 34	6	5.1	25	2	US-08-194-087-20	Sequence 20, Appli
c 35	6	5.1	25	5	PCT-US93-04648-20	Sequence 20, Appli
c 36	6	5.1	25	5	PCT-US93-04717-19	Sequence 19, Appli
c 37	6	5.1	39	4	US-09-262-773-161	Sequence 161, App
c 38	6	5.1	46	4	US-08-961-810-93	Sequence 93, Appli
c 39	6	5.1	46	4	US-08-352-902D-93	Sequence 93, Appli
c 40	6	5.1	50	1	US-08-171-389-375	Sequence 375, App
c 41	6	5.1	50	1	US-08-123-936-375	Sequence 375, App
c 42	6	5.1	50	2	US-08-475-228A-375	Sequence 375, App
c 43	6	5.1	50	3	US-08-482-080A-375	Sequence 375, App
c 44	6	5.1	50	4	US-09-354-947-375	Sequence 375, App
c 45	6	5.1	50	5	PCT-US93-12388-375	Sequence 375, App

ALIGNMENTS

RESULT 1

US-09-334-220-3

; Sequence 3, Application US/09334220

; Patent No. 6323177

; GENERAL INFORMATION:

; APPLICANT: St. Jude's Children's Research Hospital

; APPLICANT: Curran, Thomas

; TITLE OF INVENTION: INTERACTION OF REELIN WITH VERY LOW

; TITLE OF INVENTION: DENSITY LIPOPROTEIN (VLDL) RECEPTOR FOR SCREENING AND

; TITLE OF INVENTION: THERAPIES

; FILE REFERENCE: 2427/0704

; CURRENT APPLICATION NUMBER: US/09/334, 220

; CURRENT FILING DATE: 1999-06-16

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: FastSeq_for-Windows_Version 3.0

; SEQ ID NO 3

; LENGTH: 11673

; TYPE: DNA

; ORGANISM: Mus musculus

US-09-334-220-3

Alignment Scores:

Pred. No.: 1.53e-116 Length: 11673
Score: 117.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-897-438B-2 (1-117) x US-09-334-220-3 (1-11673)

Qy 1 GluGlnCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20

Db 970 GAGCAGTGGCCATCATCATGCGCAATGCTGTACCTTCTGTGAGCCGTACGGCCCT 1029

```
Qy 21 ArgGluLeuThrThrCysLeuAsnThrThrThrAlaSerValLeuGlnPheSerIle 40
Db 1030 CGAGAGTGCACACACATGCTGAACAACAACAGCATGTCTCCAGTTTCCATT 1089
Qy 41 GlySerGlySerCysArgPheSerTyrSerAspProSerIleThrValSerTyrAlaLys 60
Db 1090 GGGTCAGGATCATGCGATTAGTACTCTACCCACGATCATCTGTGTATAGCCCAAG 1149
Qy 61 AsnAsnThrAlaAspTyrIleGlnLeuGluLysIleArgAlaProSerAsnValSerThr 80
Db 1150 AACATACCGTGTGATTGATTCAGCTGGAGAAATAGAGCCCTTCCAATGTGAGCACA 1209
Qy 81 ValIleHisIleLeuTyrLeuProGluGluAlaLysGlyGluSerValGlnPheGlnTrp 100
Db 1210 GTCAATCCACATCTGTACTCTCCGAGGAGCAAGCAAGGAGCGTGCAGTTCCAGTGG 1269
Qy 101 LysGlnAspSerLeuArgValGlyGluValTyrGluAlaCysTyrAlaLeu 117
Db 1270 AACAGGACAGCGTGGAGTGGGTGAGGTGTATGAGGCGCTGCTGGGCCCTG 1320
RESULT 2
US-09-334-220-4
; Sequence 4, Application US/09334220
; Patent No. 6323177
; GENERAL INFORMATION:
; APPLICANT: St. Jude's Children's Research Hospital
; APPLICANT: Curran, Thomas
; APPLICANT: D'Arcangelo, Gabriella
; TITLE OF INVENTION: INTERACTION OF REELIN WITH VERY LOW
; TITLE OF INVENTION: DENSITY LIPOPROTEIN (VLDL) RECEPTOR FOR SCREENING AND
; TITLE OF INVENTION: THERAPIES
; FILE REFERENCE: 2427/06704
; CURRENT APPLICATION NUMBER: US/09/334,220
; CURRENT FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 11580
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-334-220-4
Alignment Scores:
Pred. No.: 2,95e-19 Length: 11580
Score: 27.00 Matches: 27
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 23.08% Indels: 0
DB: 4 Gaps: 0
US-09-897-438B-2 (1-117) x US-09-334-220-4 (1-11580)
Qy 28 LeuAsnThrThrThrAlaSerValLeuGlnPheSerIleGlySerGlySerCysArgPhe 47
Db 941 CTTAATACAAACAGCTTCTGTCTCCAAATTTTCCATTTGGTTCAGTTTCATGTCGCTTT 1000
Qy 48 SerTyrSerAspProSerIle 54
Db 1001 AGTTATTCAGACCCAGCAATC 1021
RESULT 3
US-09-894-698-14/C
; Sequence 14, Application US/09894698
; Patent No. 6469152
; GENERAL INFORMATION:
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Wisniewski, Nancy
; TITLE OF INVENTION: FLEA ALLANTOININASE NUCLEIC ACID MOLECULES, PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: FC-6-C1-C1
; CURRENT APPLICATION NUMBER: US/09/894,698
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 09/543,668
```

```
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/128,704
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 31
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Primer
US-09-894-698-14
Alignment Scores:
Pred. No.: 3.88 Length: 31
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.98% Indels: 0
DB: 4 Gaps: 0
US-09-897-438B-2 (1-117) x US-09-894-698-14 (1-31)
Qy 36 LeuGlnPheSerIleGlySer 42
Db 23 TTGCATTCAGCATAGGATCC 3
RESULT 4
US-09-302-769-28
; Sequence 28, Application US/09302769
; Patent No. 6323177
; GENERAL INFORMATION:
; APPLICANT: HILTON, Douglas J
; APPLICANT: ALEXANDER, Warren S
; APPLICANT: VINEY, Elizabeth M
; APPLICANT: WILLSON, Tracey A
; APPLICANT: RICHARDSON, Rachael T
; APPLICANT: STARR, Robyn
; APPLICANT: NICHOLSON, Sandra E
; APPLICANT: METCALF, Donald
; APPLICANT: NICOLA, Nicos A
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC AGENTS
; FILE REFERENCE: 10976Z
; CURRENT APPLICATION NUMBER: US/09/302,769
; CURRENT FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: 08/962,560
; PRIOR FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 896
; TYPE: DNA
; ORGANISM: Mouse
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (4)...(393)
; NAME/KEY: UNSURE
; LOCATION: (551)
; OTHER INFORMATION: n is unsure
; NAME/KEY: UNSURE
; LOCATION: (651)
; OTHER INFORMATION: n is unsure
US-09-302-769-28
Alignment Scores:
Pred. No.: 102 Length: 896
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.98% Indels: 0
DB: 4 Gaps: 0
```

US-09-897-438B-2 (1-117) x US-09-302-769-28 (1-896)

QY 87 LeuProGluGluAlaLysGly 93
{|||||}|||||
Db 166 CTTCTGAGGAGGCGCAAGGC 186

RESULT 5

US-09-134-001C-2718
; Sequence 2718, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2718
; LENGTH: 1125
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2718

Alignment Scores:
Pred. No.: 127 Length: 1125
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.98% Indels: 0
DB: 4 Gaps: 0

US-09-897-438B-2 (1-117) x US-09-134-001C-2718 (1-1125)

QY 83 HisleLeuTyLeuProGlu 89
{|||||}|||||
Db 248 CATATTTTATACCTGCGGAA 268

RESULT 6

US-09-222-938A-6
; Sequence 6, Application US/09222938A
; Patent No. 6437108
; GENERAL INFORMATION:
; APPLICANT: Youngman, Philip
; APPLICANT: Fritz, Chrisian
; APPLICANT: Murphy, Christopher
; APPLICANT: Guzman, Luz-Maria
; TITLE OF INVENTION: ESSENTIAL BACTERIAL GENES AND THEIR USE
; FILE REFERENCE: 07334/060001
; CURRENT APPLICATION NUMBER: US/09/222,938A
; CURRENT FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 1200
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-222-938A-6

Alignment Scores:
Pred. No.: 136 Length: 1200
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.98% Indels: 0
DB: 4 Gaps: 0

US-09-897-438B-2 (1-117) x US-09-222-938A-6 (1-1200)

QY 30 ThrThrThralaservValLeu 36
{|||||}|||||
Db 844 ACCACGACTCGTCAGTACTA 864

RESULT 7

US-08-858-207A-109
; Sequence 109, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328el Compounds
; NUMBER OF SEQUENCES: 552
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,207A
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/017670
; FILING DATE: 14-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50475
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 109:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1276 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-858-207A-109

Alignment Scores:
Pred. No.: 144 Length: 1276
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.98% Indels: 0
DB: 4 Gaps: 0

US-09-897-438B-2 (1-117) x US-08-858-207A-109 (1-1276)

QY 81 ValileHisleLeuTyLeu 87
{|||||}|||||
Db 1186 GTTATTCATATCTTACCTC 1206

RESULT 8

US-09-221-017B-933
; Sequence 933, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120


```
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1923 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 234..1811
; US-08-077-939-14

Alignment Scores:
Pred. No.: 214 Length: 1923
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.98% Indels: 0
DB: 1 Gaps: 0

US-09-897-438B-2 (1-117) x US-08-077-939-14 (1-1923)

Qy 38 PheSerIleGlySerGlySer 44
|||||
Db 958 TTTTCGATCGGTTCTGGAAGT 938

RESULT 11
US-08-461-599-14/c
; Sequence 14, Application US/08461599
; Patent No. 5658777
; GENERAL INFORMATION:
; APPLICANT: FUJII, Toshio
; APPLICANT: IWAMATSU, Akihiro
; APPLICANT: YOSHIMOTO, Hiroyuki
; APPLICANT: MINETOKI, Toshitaka
; APPLICANT: BOGAKI, Takayuki
; APPLICANT: NAGASAWA, Naoshi
; TITLE OF INVENTION: ALCOHOL ACETYLTRANSFERASE GENES AND USE
; NUMBER OF SEQUENCES: 19
; TITLE OF INVENTION: THEREOF
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,599
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/077,939
; FILING DATE: 18-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 62997/1993
; FILING DATE: 26-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 184328/1992
; FILING DATE: 18-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 49441/105 KYPA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
```

```
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1923 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 234..1811
; US-08-461-599-14

Alignment Scores:
Pred. No.: 214 Length: 1923
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.98% Indels: 0
DB: 1 Gaps: 0

US-09-897-438B-2 (1-117) x US-08-461-599-14 (1-1923)

Qy 38 PheSerIleGlySerGlySer 44
|||||
Db 958 TTTTCGATCGGTTCTGGAAGT 938

RESULT 12
US-08-461-621-14/c
; Sequence 14, Application US/08461621
; Patent No. 5686284
; GENERAL INFORMATION:
; APPLICANT: FUJII, Toshio
; APPLICANT: IWAMATSU, Akihiro
; APPLICANT: YOSHIMOTO, Hiroyuki
; APPLICANT: MINETOKI, Toshitaka
; APPLICANT: BOGAKI, Takayuki
; APPLICANT: NAGASAWA, Naoshi
; TITLE OF INVENTION: ALCOHOL ACETYLTRANSFERASE GENES AND USE
; NUMBER OF SEQUENCES: 19
; TITLE OF INVENTION: THEREOF
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,621
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/077,939
; FILING DATE: 18-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 62997/1993
; FILING DATE: 26-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 184328/1992
; FILING DATE: 18-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 49441/103 KYPA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 1923 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 234..1811
US-08-461-621-14

Alignment Scores:
Pred. No.: 214          Length: 1923
Score: 7.00           Matches: 7
Percent Similarity: 100.00%  Conservatives: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 5.98%           Indels: 0
DB: 1                  Gaps: 0

US-09-897-438B-2 (1-117) x US-08-461-621-14 (1-1923)

QY 38 PheSerIleGlySerGlySer 44
|||||
Db 958 TTTTCGATCGGTTCTGGAAGT 938

RESULT 13
US-08-465-334-14/C
; Sequence 14, Application US/08465334
; Patent No. 5728412
; GENERAL INFORMATION:
; APPLICANT: FUJII, Toshio
; APPLICANT: IWAMATSU, Akihiro
; APPLICANT: YOSHIMOTO, Hiroyuki
; APPLICANT: MINETOKI, Toshiyaka
; APPLICANT: BOGAKI, Takayuki
; APPLICANT: NAGASAWA, Naoshi
; TITLE OF INVENTION: ALCOHOL ACETYLTRANSFERASE GENES AND USE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,334
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/077,939
; FILING DATE: 18-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 62997/1993
; FILING DATE: 26-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 184328/1992
; FILING DATE: 18-JUN-1992
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 49441/101 KYPA
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1923 base pairs
; TYPE: nucleic acid
```

```
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 234..1811
US-08-465-334-14

Alignment Scores:
Pred. No.: 214          Length: 1923
Score: 7.00           Matches: 7
Percent Similarity: 100.00%  Conservatives: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 5.98%           Indels: 0
DB: 1                  Gaps: 0

US-09-897-438B-2 (1-117) x US-08-465-334-14 (1-1923)

QY 38 PheSerIleGlySerGlySer 44
|||||
Db 958 TTTTCGATCGGTTCTGGAAGT 938

RESULT 14
US-08-077-939-16/C
; Sequence 16, Application US/08077939
; Patent No. 5521088
; GENERAL INFORMATION:
; APPLICANT: FUJII, Toshio
; APPLICANT: IWAMATSU, Akihiro
; APPLICANT: YOSHIMOTO, Hiroyuki
; APPLICANT: MINETOKI, Toshiyaka
; APPLICANT: BOGAKI, Takayuki
; APPLICANT: NAGASAWA, Naoshi
; TITLE OF INVENTION: ALCOHOL ACETYLTRANSFERASE GENES AND USE
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/077,939
; FILING DATE: 18-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 184328/1992
; FILING DATE: 18-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 62997/1993
; FILING DATE: 26-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 49441/101 KYPA
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1974 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 346..1923
```


US-08-077-939-16

Alignment Scores:
Pred. No.: 220 Length: 1974
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.98% Indels: 0
DB: 1 Gaps: 0

US-09-897-438B-2 (1-117) x US-08-077-939-16 (1-1974)

Qy 38 PheSerIleGlySerGlySer 44
|||||
Db 1070 TTTTCGATCGGTTCTGGAAGT 1050

RESULT 15

US-08-461-599-16/c
: Sequence 16, Application US/08461599
: Patent No. 5658777
: GENERAL INFORMATION:
: APPLICANT: FUJII, Toshio
: APPLICANT: IWAMATSU, Akihiro
: APPLICANT: YOSHIMOTO, Hiroyuki
: APPLICANT: MINETOKI, Toshitaka
: APPLICANT: BOGAKI, Takayuki
: APPLICANT: NAGASAWA, Naoshi
: TITLE OF INVENTION: ALCOHOL ACETYLTRANSFERASE GENES AND USE
: NUMBER OF SEQUENCES: 19
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 3000 K Street, N.W., Suite 500
: CITY: Washington, D.C.
: COUNTRY: USA
: ZIP: 20007-5109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/461,599
: FILING DATE: 05-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/077,939
: FILING DATE: 18-JUN-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 62997/1993
: FILING DATE: 26-FEB-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 184328/1992
: FILING DATE: 18-JUN-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: BENT, Stephen A.
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 49441/105 KYPA
: TELEPHONE: (202)672-5300
: TELEFAX: (202)672-5399
: TELEX: 904136
: INFORMATION FOR SEQ ID NO: 16:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1974 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 346..1923
US-08-461-599-16

Alignment Scores:
Pred. No.: 220 Length: 1974
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.98% Indels: 0
DB: 1 Gaps: 0

US-09-897-438B-2 (1-117) x US-08-461-599-16 (1-1974)

Qy 38 PheSerIleGlySerGlySer 44
|||||
Db 1070 TTTTCGATCGGTTCTGGAAGT 1050

Search completed: November 6, 2002, 23:36:50
Job time : 73 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 6, 2002, 20:53:39 ; Search time 315 Seconds
(without alignments)
836.457 Million cell updates/sec

Title: US-09-897-438B-2

Perfect score: 117
Sequence: 1 EQCGTIMHGNNAVTFCEPYGP.....FQWKQDSLVRGEVVEACWAL 117

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Word size: 1

Total number of hits satisfying chosen parameters: 4368727

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q/cgn2.1/USPTO.spool/US09897438/runat_06112002_101955_11799/app_query.fasta_1.263
-DB=N_Geneseq_101002 -OFMT=fastap -SUPFIX=oligo.rng -MINMATCH=0.1 -LOOPCL=0
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR.SCORE=quality -THR.MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09897438.ecgn_1_1_79@runat_06112002_101955_11799 -NGCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	117	100.0	351	24	ABL40165
2	117	100.0	2745	24	AAD22754
3	117	100.0	11673	24	ABA92603
4	117	100.0	11673	24	AB199284
5	27	23.1	11580	24	ABN96939
6	27	23.1	11580	24	ABA92604
7	27	23.1	11632	23	ABA89484
8	24	20.5	2025	24	AAD22779
9	24	20.5	2274	24	AAD22753
10	18	15.4	413	22	AAK07720
11	18	15.4	413	22	AAK33553
12	18	15.4	413	24	ABS08422
13	14	12.0	87	22	AAK20384
14	14	12.0	87	22	AAK46490
15	14	12.0	87	24	ABS20836
16	8	6.8	1132	21	AAK81900
17	8	6.8	6407	24	ABL32943
18	8	6.8	12117	21	AAK96368
19	7	6.0	21	24	ABL40173
20	7	6.0	22	24	ABL40169
21	7	6.0	31	21	AAC93560
22	7	6.0	145	22	AAI61437
23	7	6.0	220	21	AAC08630
24	7	6.0	279	24	ABN23287
25	7	6.0	327	10	AAK90302
26	7	6.0	327	18	AAI70874
27	7	6.0	327	18	AAV15044
28	7	6.0	327	19	AAV18560
29	7	6.0	327	19	AAV18596
30	7	6.0	327	19	AAV03933
31	7	6.0	327	20	AAH22359
32	7	6.0	327	22	AAH22082
33	7	6.0	354	21	AAC05506
34	7	6.0	365	22	AAK30550
35	7	6.0	365	24	ABS05200
36	7	6.0	371	23	AAI71138
37	7	6.0	387	22	AAK40259
38	7	6.0	409	22	AAK66424
39	7	6.0	411	22	AAI01576
40	7	6.0	411	23	ABL97010
41	7	6.0	415	22	AAI19688
42	7	6.0	425	24	ABN21071
43	7	6.0	446	22	AAK75539
44	7	6.0	451	20	AAK33781
45	7	6.0	453	22	AAK59531

ALIGNMENTS

RESULT 1
ABL40165
ID ABL40165 standard; DNA; 351 BP.
XX AC ABL40165;
XX AC ABL40165;
XX AC ABL40165;
DT 21-MAY-2002 (first entry)
XX Mouse reelin protein CR-50 epitope region encoding DNA SEQ ID NO:1.
DE Mouse reelin protein CR-50 epitope region; elucidation; neuron;
XX Mouse; reelin protein CR-50 epitope region; elucidation; neuron;
KW cerebral disturbance; reelin protein; neuroprotective; gene; ds.
XX
OS Mus musculus.
XX
XX Key Location/Qualifiers
XX CDS 1..351
XX /*tag= a


```
Qy 1 GluGlnCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20
Db 970 GAGCAGTGTGGCACCACATCATGCGCAATGCTGTACCTTCTGTGAGCCGTAGGGCCCT 1029
Qy 21 ArgGluLeuThrThrCysLeuAsnThrThrThrAlaSerValLeuGlnPheSerIle 40
Db 1030 CGAGAGTGCACCACCATGCTGCTGACCAACACACAGCATGCTGCTCCAGTTTCCATT 1089
Qy 41 GlySerGlySerCysArgPheSerTyrSerAspProSerIleThrValSerTyrAlaLys 60
Db 1090 GGGTCAGGATCATGTCGATTTAGTTACTCTGACCCAGCATCACTGTGTATACGCCAAG 1149
Qy 61 AsnAsnThrAlaAspTrrIleGlnLeuGluLysIleArgAlaProSerAsnValSerThr 80
Db 1150 AACAAATACCGCTGATGGATTCAGCTGAGAAATAGAGCCCTTCCAAATGTGAGCACA 1209
Qy 81 ValIleHisIleLeuTyrLeuProGluAlaLysGlyGluSerValGlnPheGlnTrp 100
Db 1210 GTCATCCACATCCCTGTACTCTCCCGAGGAAGCCAAAGGGGAGCGGTGCAGTTCCAGTGG 1269
Qy 101 LysGlnAspSerLeuArgValGlyGluValTyrGluAlaCysTrpAlaLeu 117
Db 1270 AACAGGACAGCCCTGCGAGTGGGTGAGGTGTATGAGGCGCTGCTGGGCCCTG 1320
RESULT 3
ABA92603
ID ABA92603 standard; cDNA; 11673 BP.
XX
AC ABA92603;
XX
DT 21-MAR-2002 (first entry)
XX
DE Mouse reelin encoding cDNA SEQ ID NO:3.
XX
KW Mouse; reelin; low density lipoprotein receptor; LDLR; neuroprotective;
KW extracellular glycoprotein; nootropic; antilipaeamic; Alzheimer's disease;
KW neurodegenerative disorder; neuronal regeneration; cognitive function;
KW lipid metabolism disease; memory; developmental disorder; gene; ss.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 283..10668
FT /tag= a
FT /product= "mouse reelin"
XX
XX US6323177-B1.
XX
XX 27-NOV-2001.
XX
XX 16-JUN-1999; 99US-0334220.
XX
XX 16-JUN-1999; 99US-0334220.
XX
XX (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
XX Curran T, D'Arcangelo G;
XX
XX WPI; 2002-096596/13.
XX
XX P-PSDB; ABB05008.
XX
XX Novel composition useful for screening compounds that modulate Reelin
XX binding to low density lipoprotein receptor, comprising an isolated
XX Reelin polypeptide and low density lipoprotein receptor .
XX
XX Example 1; Column 65-74; 45pp; English.
XX
XX The present invention describes a composition (I) comprising an
XX isolated reelin protein (II) bound to an isolated low density lipoprotein
XX receptor (LDLR) (III). (II) is an extracellular glycoprotein of
XX approximately 385 kDa containing a small region of similarity with
XX F-spondin at the N terminus, a stretch of positively charged amino
```

```
CC acids at the C terminus, and a series of eight internal repeats of
CC 330-390 amino acids, each repeat containing two related sub-domains
CC that flank a pattern of conserved cysteine residues known as an
CC epidermal growth factor (EGF)-like motif. (I) has neuroprotective,
CC nootropic and antilipaeamic activities, and can be used as a modulator
CC of reelin-LDLR interaction. (I) is useful in screen for compounds that
CC modulate reelin binding to an LDLR, in an assay system, where the assay
CC system comprises a microplate array and an automated robotic
CC microprocessor controlled system for adding and removing reagents to
CC the microplate array. The compounds identified by the above screening
CC method are useful as therapeutic agents to provide or alleviate a
CC diverse spectrum of diseases including neurodegenerative disorders such
CC as Alzheimer's disease, to facilitate neuronal regeneration after
CC injury, to prevent or alleviate lipid metabolism diseases, to enhance
CC cognitive functions and memory or to ameliorate other developmental
CC disorders. The present sequence encodes mouse (Mus musculus) reelin,
CC which is used in the exemplification of the present invention.
```

XX Sequence 11673 BP; 2831 A; 2985 C; 2985 G; 2872 T; 0 other;

Alignment Scores:

Pred. No.: 2.7e-110 Length: 11673
Score: 117.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

US-09-897-438B-2 (1-117) x ABA92603 (1-11673)

```
Qy 1 GluGlnCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20
Db 970 GAGCAGTGTGGCACCACATCATGCGCAATGCTGTACCTTCTGTGAGCCGTAGGGCCCT 1029
Qy 21 ArgGluLeuThrThrCysLeuAsnThrThrThrAlaSerValLeuGlnPheSerIle 40
Db 1030 CGAGAGTGCACCACCATGCTGACCAACACACAGCATCTGCTCCAGTTTCCATT 1089
Qy 41 GlySerGlySerCysArgPheSerTyrSerAspProSerIleThrValSerTyrAlaLys 60
Db 1090 GGGTCAGGATCATGTCGATTTAGTTACTCTGACCCAGCATCACTGTGTATACGCCAAG 1149
Qy 61 AsnAsnThrAlaAspTrrIleGlnLeuGluLysIleArgAlaProSerAsnValSerThr 80
Db 1150 AACAAATACCGCTGATGGATTCAGCTGAGAAATAGAGCCCTTCCAAATGTGAGCACA 1209
Qy 81 ValIleHisIleLeuTyrLeuProGluAlaLysGlyGluSerValGlnPheGlnTrp 100
Db 1210 GTCATCCACATCCCTGTACTCTCCCGAGGAAGCCAAAGGGGAGCGGTGCAGTTCCAGTGG 1269
Qy 101 LysGlnAspSerLeuArgValGlyGluValTyrGluAlaCysTrpAlaLeu 117
Db 1270 AACAGGACAGCCCTGCGAGTGGGTGAGGTGTATGAGGCGCTGCTGGGCCCTG 1320
```

RESULT 4

ABI99284
ID ABI99284 standard; cDNA; 11673 BP.

XX

AC ABI99284;

XX

DT 07-MAR-2002 (first entry)

XX

DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:128.

XX

KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;

XX

KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.

XX

OS Mus musculus.

XX

PN WO2001188188-A2.

XX

PD 22-NOV-2001.

XX

```
PF 18-MAY-2001; 2001WO-JP04192.
XX
PR 18-MAY-2000; 2000JP-0145977.
XX
PA (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX
DR WPI; 2002-034733/04.
DR P-PSDB; ABB57065.
XX
XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or
PT by determining the expression profile of a gene group comprising these
PT genes -
XX
PS Claim 2; Page 362-385; 2690pp; English.
XX
CC The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC expression levels of particular genes (ABI99202 to ABI99912, encoding
CC the protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The
CC expression levels or expression profiles produced by these genes are
CC used as an indicator when screening for ischaemic condition-improving
CC drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914
CC represent PCR primers for a mouse ischaemic condition related sequence,
CC which are used in the exemplification of the present invention.
XX
SQ Sequence 11673 BP; 2831 A; 2985 C; 2985 G; 2872 T; 0 other;

Alignment Scores:
Pred. No.: 2,7e-110 Length: 11673
Score: 117.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

US-09-897-438B-2 (1-117) x ABI99284 (1-11673)
QY 1 GluGlnCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20
DB 970 GAGCAGTGTGGCACCACATCATGTCATGGCAATGCTGTCACCTTCTGTGAGCGGTACGGCCT 1029
QY 21 ArgGluLeuThrThrThrCysLeuAsnThrThrThrAlaSerValLeuGlnPheSerIle 40
DB 1030 CGAGAGTGACCCACCATGCTGTAACACACACACATGCTGCTCCAGTTTTCATT 1089
QY 41 GlySerGlySerCysArgPheSerTyrSerAspProSerIleThrValSerTyrAlaLys 60
DB 1090 GGGTCAGGATCATGTCGATTAGTTACTGACCCAGCATCATCTGTCATAGCCCAAG 1149
QY 61 AsnAsnThrAlaAspTrpIleGlnLeuGluLysIleArgAlaProSerAsnValSerThr 80
DB 1150 AACAAATACCCGCTGATTCGATTCAGCTGGAGAAAATTAGAGCCCTTCCAAATGTGAGCACA 1209
QY 81 ValIleHisIleLeuTyrLeuProGluGluAlaLysGlyGluSerValGlnPheGlnTrp 100
DB 1210 GTCATCCACATCTGTAACCTCCCGAGGAGAGCCAAAGGGAGAGCGGTGACAGTTCCAGTGG 1269
QY 101 LysGlnAspSerLeuArgValGlyGluValTyrGluAlaCysTrpAlaLeu 117
DB 1270 AAACAGCAGACGCTCGCAGTGGGTGAGTGTATGAGGCGCTGCTGGGCCCTG 1320

RESULT 5
ABN96939
ID ABN96939 standard; DNA; 11580 BP.
XX
```

```
AC ABN96939;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #3437 used to diagnose liver cancer.
XX
KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumour; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
XX
PN WO200229103-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US30589.
XX
PR 02-OCT-2000; 2000US-237054P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX
DR WPI; 2002-426119/45.
XX
XX Diagnosing and detecting the progression of liver cancer,
PT hepatocellular carcinoma or metastatic liver tumor in a patient, in a
PT involves detecting the level of expression of two or more genes, in a
PT liver tissue sample -
XX
PS Claim 1; SEQ ID NO 3437; 298pp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 11580 BP; 3014 A; 2696 C; 2753 G; 3116 T; 1 other;

Alignment Scores:
Pred. No.: 2,16e-17 Length: 11580
Score: 27.00 Matches: 27
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 23.08% Indels: 0
DB: 24 Gaps: 0

US-09-897-438B-2 (1-117) x ABN96939 (1-11580)
QY 28 LeuAsnThrThrThrAlaSerValLeuGlnPheSerIleGlySerGlySerCysArgPhe 47
DB 941 CTTAATACACAAACAGCTTCTGTCTCCCAATTTTCCATTTGGTCAGGTTTCATGTCGCTTT 1000
QY 48 SerTyrSerAspProSerIle 54
DB 1001 AGTATTTCAGACCCAGCATC 1021

RESULT 6
ABN92604
ID ABA92604 standard; cDNA; 11580 BP.
XX
```

AC ABA92604;

DT 21-MAR-2002 (first entry)

DE Human reelin encoding cDNA SEQ ID NO:4.

XX

KW Human; reelin; low density lipoprotein receptor; LDLR; neuroprotective;

KW extracellular glycoprotein; nontropic; antilipemic; Alzheimer's disease;

KW neurodegenerative disorder; neuronal regeneration; cognitive function;

KW lipid metabolism disease; memory; developmental disorder; gene; ss.

XX

OS Homo sapiens.

XX

PH Key Location/Qualifiers

FT 176..10558

FT /*tag= a

FT /product= "human reelin"

XX

PN US6323177-B1.

XX

XX 27-NOV-2001.

PD

XX 16-JUN-1999; 99US-0334220.

XX

XX 16-JUN-1999; 99US-0334220.

PR

XX (SUJUD-) ST JUDE CHILDREN'S RES HOSPITAL.

PA

XX Curran T, D'Arcangelo G;

PI

XX WPI: 2002-096596/13.

DR

DR P-PSDB; ABB05007.

XX

XX Novel composition useful for screening compounds that modulate Reelin

PT binding to low density lipoprotein receptor, comprising an isolated

PT Reelin polypeptide and low density lipoprotein receptor

XX

XX Example 1; Column 75-84; 45pp; English.

PS

XX

CC The present invention describes a composition (I) comprising an

CC isolated reelin protein (III) bound to an isolated low density lipoprotein

CC receptor (LDLR) (III). (II) is an extracellular glycoprotein of

CC approximately 385 kDa containing a small region of similarity with

CC F-spondin at the N terminus, a stretch of positively charged amino

CC acids at the C terminus, and a series of eight internal repeats of

CC 350-390 amino acids, each repeat containing two related sub-domains

CC that flank a pattern of conserved cysteine residues known as an

CC epidermal growth factor (EGF)-like motif. (I) has neuroprotective,

CC nontropic and antilipemic activities, and can be used as a modulator

CC of reelin-LDLR interaction. (I) is useful in screen for compounds that

CC modulate reelin binding to an LDLR, in an assay system, where the assay

CC system comprises a microplate array and an automated robotic

CC microprocessor controlled system for adding and removing reagents to

CC the microplate array. The compounds identified by the above screening

CC method are useful as therapeutic agents to provide or alleviate a

CC diverse spectrum of diseases including neurodegenerative disorders such

CC as Alzheimer's disease, to facilitate neuronal regeneration after

CC injury, to prevent or alleviate lipid metabolism diseases, to enhance

CC cognitive functions and memory or to ameliorate other developmental

CC disorders. The present sequence encodes human reelin, which is used in

CC the exemplification of the present invention.

XX

SQ Sequence 11580 BP; 3014 A; 2696 C; 2753 G; 3116 T; 1 other;

Alignment Scores:

Pred. No.:	2,16e-17	Length:	11580
Score:	27.00	Matches:	27
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	23.08%	Indels:	0
DB:	24	Gaps:	0

US-09-897-438B-2 (1-117) x ABA92604 (1-11580)

Qy 28 LeuAsnThrThrThrAlaSerValLeuGlnPheSerIleGlySerGlySerCysArgPhe 47

Db 941 CTTAATAACAACACAGCTTCTGTCTCCTCAATTTTCCATTTGGTTCAGTTCATGTCGCTTT 1000

Qy 48 SerTyrSerAspProSerIle 54

Db 1001 AGTTATTACAGACCCCGACATC 1021

RESULT 7

AA89484

ID AAS89484 standard; cDNA; 11632 BP.

XX

AC AAS89484;

XX

DT 13-FEB-2002 (first entry)

XX

DE DNA encoding novel human diagnostic protein #25288.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX

OS Homo sapiens.

XX

XX WO200175067-A2.

PN

XX 11-OCT-2001.

PD

XX 30-MAR-2001; 2001WO-US08631.

PF

XX 31-MAR-2000; 2000US-0540217.

PR

XX 23-AUG-2000; 2000US-0649167.

PR

XX (HYSE-) HYSEQ INC.

PA

XX Drmanac RT, Liu C, Tang YT;

PI

XX WPI: 2001-639362/73.

DR

DR P-PSDB; ABG25297.

XX

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity

XX

XX Claim 1; SEQ ID NO 25288; 103pp; English.

PS

XX The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving

CC (II). (II) is useful for generating antibodies against it, detecting or

CC quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical

CC imaging of sites expressing (II). (I) and (II) are useful for treating

CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. AAS64197-AAS94564 represent novel human

CC diagnostic coding sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 11632 BP; 3028 A; 2716 C; 2770 G; 3117 T; 1 other;

Alignment Scores:

Pred. No.:	2.17e-17	Length:	11632
------------	----------	---------	-------

Score: 27.00 Matches: 27
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 23.08% Indels: 0
DB: 23 Gaps: 0

US-09-897-438B-2 (1-117) x AAS89484 (1-11632)

Qy 28 LeuAsnThrThrAlaSerValLeuGlnPheSerIleGlySerGlySerCysArgPhe 47
|||||
Db 941 CTTAATACAACACAGCTTCGTCTCCCAATTTCCATTGGTCAGGTCATGTCGCTTT 1000
|||||

Qy 48 SerTyrSerAspProSerIle 54
|||||
Db 1001 AGTTATTCAGACCCAGCATC 1021
|||||

RESULT 8
AAD22779
ID AAD22779 standard; cDNA; 2025 BP.
XX
AC AAD22779;
XX
DT 26-FEB-2002 (first entry)
XX
DE Xenopus laevis intact reelin cDNA.
XX
KW Frog; reelin; F-spondin domain; CR-50 epitope; gene therapy; agyria;
KW polymicrogyria; ectopic gray matter; ss.
XX
OS Xenopus laevis.
XX
FH Key Location/Qualifiers
FT CDS 157..2025
FT /*tag= a
FT /product= "Xenopus intact reelin protein"
FT /trans_except= (pos:1606..1608, aa:Cys)
FT /transl_except= (pos:1639..1641, aa:Cys)
FT /note= "CDS does not include stop codon"
FT /partial
XX
PN EP1149844-A2.
XX
XX 31-OCT-2001.
XX
PF 11-APR-2001; 2001EP-0303411.
XX
PR 11-APR-2000; 2000JP-0109954.
XX
PA (RIKE) RIKEN KK.
XX
PI Mikoshiba K, Tabata H, Nakajima K;
XX
XX WPI; 2002-019320/03.
XX
XX P-PSDB; AAEL3607.
XX
PT Novel truncated Reelin protein containing F-spondin domain and CR-50
PT recognition site of Reelin protein, but not having Reelin repeat site,
PT useful to treat diseases including agyria due to abnormal neuron
PT alignment
XX
XX
PS Example 1; Fig 1A; 47pp; English.
XX
XX The invention relates to a truncated Reelin protein comprising a
XX F-spondin domain and a CR-50 recognition site but no reelin protein
XX repeat site. Reelin is an essential molecule in developing a normal
XX laminated structure of cerebrum. The truncated reelin protein and its
XX DNA are useful for treating diseases including agyria, polymicrogyria,
XX and ectopic gray matter due to abnormal neuronal alignment. Truncated
XX reelin protein DNA is useful in gene therapy. The present sequence is
XX a cDNA encoding Xenopus laevis intact reelin protein.
XX
SQ Sequence 2025 BP; 541 A; 419 C; 460 G; 604 T; 1 other;
XX

Alignment Scores:
Pred. No.: 5.13e-15 Length: 2025
Score: 24.00 Matches: 24
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.51% Indels: 0
DB: 24 Gaps: 0

US-09-897-438B-2 (1-117) x AAD22779 (1-2025)

Qy 29 AsnThrThrThrAlaSerValLeuGlnPheSerIleGlySerGlySerCysArgPheSer 48
|||||
Db 931 AACACAACACTAGGCATCTGTTTTCAGTTTCTATTGGTCAGATCGTCAGGTTTCAGC 990
|||||

Qy 49 TyrSerAspPro 52
|||||
Db 991 TATTCAGACCCCT 1002
|||||

RESULT 9
AAD22753
ID AAD22753 standard; cDNA; 2274 BP.
XX
AC AAD22753;
XX
DT 26-FEB-2002 (first entry)
XX
DE Xenopus laevis truncated reelin cDNA.
XX
KW Frog; reelin; F-spondin domain; CR-50 epitope; gene therapy; agyria;
KW polymicrogyria; ectopic gray matter; ss.
XX
OS Xenopus laevis.
XX
FH Key Location/Qualifiers
FT CDS 157..1455
FT /*tag= a
FT /product= "X. laevis truncated reelin protein"
FT sig_peptide 157..234
FT /*tag= b
FT mat_peptide 235..1452
FT /*tag= c
FT misc_feature 241..726
FT /*tag= d
FT misc_feature 847..1197
FT /*tag= e
FT polyA_signal 2254..2259
FT /*tag= f
XX
XX EP1149844-A2.
XX
XX 31-OCT-2001.
XX
PF 11-APR-2001; 2001EP-0303411.
XX
PR 11-APR-2000; 2000JP-0109954.
XX
PA (RIKE) RIKEN KK.
XX
XX Mikoshiba K, Tabata H, Nakajima K;
XX
XX WPI; 2002-019320/03.
XX
XX P-PSDB; AAEL3605.
XX
PT Novel truncated Reelin protein containing F-spondin domain and CR-50
PT recognition site of Reelin protein, but not having Reelin repeat site,
PT useful to treat diseases including agyria due to abnormal neuron
PT alignment
XX
XX
PS Claim 8; Page 11-16; 47pp; English.
XX
XX

CC The invention relates to a truncated reelin protein comprising a
CC F-spondin domain and a Cr-50 recognition site but no reelin protein
CC repeat site. Reelin is an essential molecule in developing a normal
CC laminated structure of cerebrum. The truncated reelin protein and its
CC DNA are useful for treating diseases including agyria, polymicrogyria,
CC and ectopic gray matter due to abnormal neuronal alignment. Truncated
CC reelin protein DNA is useful in gene therapy. The present sequence is
CC a cDNA encoding *Xenopus laevis* truncated reelin protein.

XX Sequence 2274 BP; 660 A; 450 C; 469 G; 694 T; 1 other;

Alignment Scores:
Pred. No.: 5.73e-15 Length: 2274
Score: 24.00 Matches: 24
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.51% Indels: 0
DB: 24 Gaps: 0

US-09-897-438B-2 (1-117) x AAD22753 (1-2274)

QY 29 AsnThrThrAlaSerValLeuGlnPheSerIleGlySerGlySerCysArgPheSer 48
|||||
Db 931 AACACAACTACGGCATCTGTTTGCAGTTTCTATTGGTTCAGGATCGTCAGGTTTCAGC 990

QY 49 TyrSerAspPro 52
|||||

Db 991 TATTGAGACCCCT 1002

RESULT 10

AAK07720/C

ID AAK07720 standard; DNA; 413 BP.

XX AAK07720;

XX 05-NOV-2001 (first entry)

XX Human brain expressed single exon probe SEQ ID NO: 7711.
XX Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer; ss.

XX Homo sapiens.

XX WO200157275-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00667.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX Example 4; SEQ ID NO: 7711; 650pp + Sequence Listing; English.
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,

CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention.

XX Sequence 413 BP; 134 A; 86 C; 67 G; 126 T; 0 other;

Alignment Scores:
Pred. No.: 1.76e-09 Length: 413
Score: 18.00 Matches: 18
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 15.38% Indels: 0
DB: 22 Gaps: 0

US-09-897-438B-2 (1-117) x AAK07720 (1-413)

QY 6 IleMethHisGlyAsnAlaValThrPheCysGluProTyrGlyProArgGluLeu 23
|||||

Db 308 ATTATGCATGGCAATGCCGTCACCTTCTGTGACCATATGCCACGAGAACTG 255

RESULT 11

AAK33553/C

ID AAK33553 standard; DNA; 413 BP.

XX AAK33553;

XX 06-NOV-2001 (first entry)

XX Human bone marrow expressed single exon probe SEQ ID NO: 8110.
XX Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.

XX Homo sapiens.

XX WO200157276-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00668.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -
XX Example 4; SEQ ID NO: 8110; 658pp + Sequence Listing; English.
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukaemia and myeloma. The present sequence is one of
XX the probes of the invention.

XX Sequence 413 BP; 134 A; 86 C; 67 G; 126 T; 0 other;

Alignment Scores:
Pred. No.: 1.76e-09 Length: 413
Score: 18.00 Matches: 18

```
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 15.38%              Indels: 0
DB: 22                           Gaps: 0

US-09-897-438B-2 (1-117) x AAK33553 (1-413)

QY 6 IleMethHisGlyAsnAlaValThrPheCysGluProTyrGlyProArgGluLeu 23
   |||||||
Db 308 ATTATGATGCAATGCCGTCTCTGTGAACCATATGCGCCACGAGAACTG 255

RESULT 12
ABS08422/c
ID ABS08422 standard; DNA; 413 BP.
XX
AC ABS08422;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe from lung SEQ ID No 8413.
XX
KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemorrhoidosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US00665.
XX
PR 04-FEB-2000; 2000US-180312P.
PR 26-MAY-2000; 2000US-207456P.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-234687P.
PR 27-SEP-2000; 2000US-236359P.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples -
XX
PS Claim 1; SEQ ID No 8413; 634pp; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of
CC probes; the novel set of probes which hybridise at high stringency to a
CC nucleic acid expressed in the human lung; measuring gene expression in a
CC sample derived from human lung, comprising (a) contacting the array with
CC a collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of
CC the array; identifying exons in a eukaryotic genome, comprising
CC (a) algorithmically predicting at least one exon from genomic sequences
CC of the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
```

```
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC using human lung derived mRNA and for the study of lung diseases
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC haemorrhoidosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
CC and hyaline membrane disease. The present sequence is a single exon
CC probe of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 413 BP; 134 A; 86 C; 67 G; 126 T; 0 other;

Alignment Scores:
Pred. No.: 1.76e-09      Length: 413
Score: 18.00            Matches: 18
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 15.38%              Indels: 0
DB: 24                    Gaps: 0

US-09-897-438B-2 (1-117) x ABS08422 (1-413)

QY 6 IleMethHisGlyAsnAlaValThrPheCysGluProTyrGlyProArgGluLeu 23
   |||||||
Db 308 ATTATGATGCAATGCCGTCTCTGTGAACCATATGCGCCACGAGAACTG 255

RESULT 13
AAK20384/c
ID AAK20384 standard; DNA; 87 BP.
XX
AC AAK20384;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 20375.
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
```

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
XX Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
XX Example 4; SEQ ID NO: 20375; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention.
XX
XX Sequence 87 BP; 21 A; 25 C; 19 G; 22 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 5.38e-06 Length: 87
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.97% Indels: 0
DB: 22 Gaps: 0
US-09-897-438B-2 (1-117) x AAK46490 (1-87)
Qy 6 IleMethHisGlyAsnAlaValThrPheCysGluProTyrGly 19
Db 44 ATTATGTCATGCGCATGCGCTCACCTTCTGTGAACCATATGGC 3
RESULT 14
ID AAK46490/c
AC AAK46490;
XX
XX 06-NOV-2001 (first entry)
XX Human bone marrow expressed single exon probe SEQ ID NO: 21047.
XX
XX Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
XX Homo sapiens.
XX
XX WO200157276-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00668.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 21047; 658pp + Sequence Listing; English.
PS

XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention.
XX
XX Sequence 87 BP; 21 A; 25 C; 19 G; 22 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 5.38e-06 Length: 87
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.97% Indels: 0
DB: 22 Gaps: 0
US-09-897-438B-2 (1-117) x AAK46490 (1-87)
Qy 6 IleMethHisGlyAsnAlaValThrPheCysGluProTyrGly 19
Db 44 ATTATGTCATGCGCATGCGCTCACCTTCTGTGAACCATATGGC 3
RESULT 15
ID ABS20836/c
AC ABS20836;
XX
XX 19-AUG-2002 (first entry)
XX
XX Human genome-derived single exon probe ORF from lung SEQ ID NO 20827.
XX
XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease; open reading frame; ORF.
XX
XX Homo sapiens.
XX
XX WO200186003-A2.
XX
XX 15-NOV-2001.
XX
XX 30-JAN-2001; 2001WO-US00665.
XX
XX 04-FEB-2000; 2000US-180312P.
XX 26-MAY-2000; 2000US-207456P.
XX 30-JUN-2000; 2000US-0608408.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-234687P.
XX 27-SEP-2000; 2000US-236359P.
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2002-114183/15.
XX
XX Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples -
XX
XX Claim 4; SEQ ID NO 20827; 634pp; English.
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived

CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of
CC probes; the novel set of probes which hybridise at high stringency to a
CC nucleic acid expressed in the human lung; measuring gene expression in a
CC sample derived from human lung, comprising (a) contacting the array with
CC a collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of
CC the array; identifying exons in a eukaryotic genome, comprising
CC (a) algorithmically predicting at least one exon from genomic sequences
CC of labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC using human lung derived mRNA and for the study of lung diseases
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
CC and hyaline membrane disease. The present sequence is a single exon
CC probe open reading frame of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 87 BP; 21 A; 25 C; 19 G; 22 T; 0 other;

Alignment Scores:
Pred. No.: 5.38e-06 Length: 87
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.97% Indels: 0
DB: 24 Gaps: 0

US-09-897-438B-2 (1-117) x ABS20836 (1-87)

Qy 6 IleMethHisGlyAsnAlaValThrPheCysGluProTyrGly 19
Db 44 ATTATGATGCGCAATGCGTCACCTTCTGTGACCATATGCC 3

Search completed: November 6, 2002, 22:12:10
Job time : 324 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 6, 2002, 20:55:05 ; Search time 2763 Seconds
(without alignments)
1232.366 Million cell updates/sec

Title: US-09-897-438B-2
Perfect score: 117

Sequence: 1 EQCGTTHGNNAVTFCEPYGP.....FQWKQDSLVRGEVVEACWAL 117

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Word size: 1

Total number of hits satisfying chosen parameters: 4106490

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool/US09897438/runat_06112002_101956_11805/app_query.fasta_1.263
-DB=GenEmbl -QFMT=fastap -SUFFIX=oligo.rge -MINMATCH=0.1 -LOOPCI=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LISF=45
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptp
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09897438 -CGN_1_1_2659_@runat_06112002_101956_11805 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_em.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_mu.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rnd.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	117	100.0	11673	6	AX305377 Sequence
2	117	100.0	11673	10	U24703 Mus musculus
c 3	45	38.5	163985	10	AC121878 Mus musculus
5	45	38.5	183156	2	AC023062 Mus musculus
6	27	23.1	11187	10	AB049473 Rattus no
7	27	23.1	11580	6	AX410790 Sequence
8	27	23.1	11580	9	U79716 Human reeli
9	27	23.1	11580	11	G30936 SWSS2926 Er
c 10	23	23.1	11580	11	G30938 SWSS3176 Er
11	22	18.8	126130	2	AC124933 Rattus no
12	22	18.8	202764	2	AC095877 Rattus no
13	18	15.4	10634	5	AF090441 Gallus ga
c 14	18	15.4	93163	9	AC001121 Human BAC
c 15	18	15.4	185996	2	AC041023 Homo sapi
16	11	9.4	158	4	AF232904 Bos tauru
17	10	8.5	49858	2	AC008252 Drosophil
c 18	10	8.5	55700	2	AC017867 Drosophil
19	10	8.5	196594	3	AC010564 Drosophil
20	10	8.5	300412	3	AF003471 Drosophil
c 21	9	7.7	150869	2	AF005550 Oryza sat
c 22	9	7.7	153805	2	AP004661 Oryza sat
23	9	7.7	176801	2	AP005493 Oryza sat
c 24	9	7.7	177469	2	AC128345 Rattus no
c 25	9	7.7	216589	2	AC091741 Homo sapi
c 26	9	7.7	217488	9	AP003486 Homo sapi
27	8	6.8	438	11	G02057 STScl9B12 L
28	8	6.8	567	9	AB018508 Homo sapi
c 29	8	6.8	1132	6	AX040663 Sequence
c 30	8	6.8	1540	8	AY128338 Arabidops
c 31	8	6.8	1765	8	AY120722 Arabidops
c 32	8	6.8	1861	8	YSPPTC1X Schizosacch
c 33	8	6.8	2378	8	AF050673 Cossypium
c 34	8	6.8	2632	5	GGCTR3 X13753 Chicken mRN
c 35	8	6.8	2926	8	AF173640 Arabidops
c 36	8	6.8	3361	9	AF004169 Pan trogl
c 37	8	6.8	4146	8	AF001035 Arabidops
c 38	8	6.8	4338	1	PVU60522 Paracoccus
39	8	6.8	4514	10	AB048843 Mus muscu
c 40	8	6.8	4652	2	AC015122 Drosophil
c 41	8	6.8	6007	8	SCYPL274W S.cerevisia
c 42	8	6.8	6069	8	SCYPL276W S.cerevisia
c 43	8	6.8	6407	6	AX345845 Sequence
c 44	8	6.8	10378	8	AX117389 S.pombe c
c 45	8	6.8	30544	9	ALI139108 Human DNA

ALIGNMENTS

RESULT 1

```
AX305377
LOCUS AX305377 11673 bp DNA linear PAT 11-DEC-2001
DEFINITION Sequence 128 from Patent WO0188188.
ACCESSION AX305377
VERSION AX305377.1 GI:17644926
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE
AUTHORS Ishikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y.
TITLE Method for examining ischemic conditions
JOURNAL Patent: WO 0188188-A 128 22-NOV-2001;
SCHOOL Juridical Person Nihon University (JP)
FEATURES
Location/Qualifiers
1..11673
/organism="Mus musculus"
/db_xref="taxon:10090"
BASE COUNT 2831 a 2985 c 2985 g 2872 t
ORIGIN
Alignment Scores:
Pred. No.: 6.56e-116 Length: 11673
Score: 117.00 Matches: 117
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-897-438B-2 (1-117) x AX305377 (1-11673)
Qy 1 GluGlnCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20
|||||
Db 970 GAGCAGTGTGCACCATCATGCATGGAATGCTGTCACCTCTCTGTGAGCGGTACGGCCCT 1029
Qy 21 ArgGluLeuThrThrThrCysLeuAsnThrThrThrAlaSerValLeuGlnPheSerIle 40
|||||
Db 1030 CGAGAGCTGACACCATCATGCTGCTGAACACAAACACATGTCCTCCATTTTCCATT 1089
Qy 41 GlySerGlySerCysArgPheSerTyrSerAspProSerIleThrValSerTyrAlaLys 60
|||||
Db 1090 GGGTCAGGATCATGTCGATTTAGTACTCTGACCCAGCATCATCTGTCTACGCCAAG 1149
Qy 61 AsnAsnThrAlaAspTrpIleGlnLeuGluLysIleArgAlaProSerAsnValSerThr 80
|||||
Db 1150 AACAAATACCGTGTGATGATTCAGTTCAGTGCAGAAATAGAGCCCTTCCAATGTGAGCACA 1209
Qy 81 ValIleHisIleLeuTyrLeuProGluGluAlaLysGlyGluSerValGlnPheGlnTrp 100
|||||
Db 1210 GTCATCCACATCCTGTACCTCCCGAGGAGCCAAAGGAGAGCGTGCAGTTCACGTGG 1269
Qy 101 LysGlnAspSerLeuArgValGlyGluValTyrGluAlaCysTrpAlaLeu 117
|||||
Db 1270 AACAGAGACAGCCCTGCAGTGGGTGAGTGTATGAGGCGCTGCTGGGCCCTG 1320
RESULT 2
MMU24703
LOCUS MMU24703 11673 bp mRNA linear ROD 19-DEC-1997
DEFINITION Mus musculus reelin mRNA, complete cds.
ACCESSION U24703
VERSION U24703.1 GI:2702252
KEYWORDS
SOURCE
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE
AUTHORS D'Arcangelo,G., Miao,G.G., Chen,S.C., Soares,H.D., Morgan,J.I. and Curran,T.
TITLE A protein related to extracellular matrix proteins deleted in the mouse mutant reeler
JOURNAL Nature 374 (6524), 719-723 (1995)
```

```
95231649
MEDLINE
PUBMED 7715726
REFERENCE 2 (bases 1 to 11673)
AUTHORS D'Arcangelo,G.
TITLE Direct Submmission
JOURNAL Submitted (10-APR-1995) Roche Institute of Molecular Biology, 340
Kingsland St., Nutley, NJ 07110, USA
3 (bases 1 to 11673)
D'Arcangelo,G.
Direct Submmission
Submitted (19-DEC-1997) Dev. Neurobiology, St. Jude Childr. Res.
Hosp, 332 N. Lauderdale, Memphis, TN 38105, USA
REMARK Sequence update by submitter
COMMENT On Dec 19, 1997 this sequence version replaced gi:902486.
FEATURES
Location/Qualifiers
1..11673
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="5"
/map="D5Gmrl"
/tissue_type="cerebellum"
/dev_stage="adult"
1..282
283..10668
/codon_start=1
/product="reelin"
/protein_id="AAB91599.1"
/db_xref="GI:2702253"
/translation="MERGCPAPRALVLAVLLLLATLARAATGYPRFSPFFFLCTHH
GELEGGEGEVLISHIAGNTIYYPGGYHVTITSTFFDGLLVLTGLVTSIQSS
QSTGGSAFGIMSDHQFNQPCSVASHVSHLPTTNLSFIWAPAGTGCVFNFA
TATHRGQVIFKDALAQOLCQEQATATAYSHLAEIHSDSVILRDQFDSYQLELNFN
IWEPCNCEMGEOCTIMHNAVTFCEPRLTNTCLNTTASVLOSIGSGSCRF
SYSDSITVSYAKNTADMTQLEKIRAPSNVSTIHLXLPPEAKESVQFQKQSL
RVEGVACWALDNILVINSAREVLENDLPDVTGNLWLPFCATVVKHSCQSDGNSI
YFHNGESFNFAITRDVLDSTEDIQESESFEFSQPTGMDILGAVVGDGCTVESGL
SLVFLKDKRGKJCTPYMDTGYGNLRFYFMVGGIDGPGVSHENDIILYAKIEGKEHI
ADLTLYSSYKPSLVSVINPELOTPATKCLRQKSHOQYNNRVWADFFHLPVLP
TSMHMIQFSINLGCCTHQPNSVLEFSTNHRGSSLLHTECLPETAAGPHLPSTV
YSENYSWNRTITPLPNAALTRDIRWRQTGPIILGNWAIIDNVIIGPSCLAFCSGR
GOCNTRHGCDCDFGSPACEMASQTPMFISEFSGSARLSYHNFYSIRGAESVSCGG
VLASGALVFNKGRQLITSLDSSSRFLQTLRLGSKSVLSTRAPDQPGEGVLL
HYSYDNGITWKLHYSYVNYHEPRIISVELPDARQFGIQFRMWPYHSSQGEDVMA
IDEIVMTSLFNFSISLDFTNLVEVTSGLFNLGNQVPCYCHDWTLCFTGDKSLASSR
YVBTQSMQIGASYMIQFSLVMGCGQKYPHMDNQVKLEYSAHGLTHLWVQEECLPSM
PSCOEFTSASLYHASEFTQWRVTVLPKPTWSGATRFERMSQSYTYAODPWLADNLYI
GOOCNMGCSHGSCDHGCGCDOGYOCTECHPEALPSTIMSDENPSSWESDQVEI
GGEVYPECGCGVYVSSGSLYFSKAGRQLVSWDLDTSWDFVQFTIQIGESAACNK
PDSREGLILQYSNNGGIOMHLLAEMTFSDFSKPRFYLELPAAGTKTCTFRFRMKPV
FGEDYDQWAVDDIILSEKQKQVIPVNPFTLPQNYEKPAFDYPMQNSVWLMLANE
GMAKNSDFCATTPSAMVFGKSDGDFEAVTRDLTKPGYVLQNLINGCTSQFSTAPV
LLQYSHDAGMSWFLKKEGCFPASAAGCEGNSRELSEPTVYTYGTDFEENRTITAIPT
SLASSKTRFRWTOESSOKNVPFGLDGYVISEPCPSYCSGHGDCISGVCFCDLGTA
AOGTCVSNTPNISEMDFREGKLSPLWYKITGGVGTGCTLNDGSLRYFLNGLKREA
RTVPLDTRNISLVQFYIQIGSKTSGITYITPRARYEGLVVQYNSDNLGHLHLLRELD
MSFLEPIIISIDLPREAKPATAFRWPQHGKSAQWALGDVLIIGNVSSQTFQDK
LDGSDILQANWYRIQGGQVDICLSMDTALIFTENIGNPRYAEFTDFHYSESFLQWE
NMNGCSKPSGAGHIOLOYSLNNKDWLTBECVPTTGYCHVTYESSVTSRFFOWM
RVTVYPLATNSPRTFRWIOINTVAGADSWAIDNVILASGCPMWSGRCIGDSRC
VCDRGFGGFCYVPVPLPSLKDFFNGNLHPDLWPEVYGAERGNLNGETIKSTCLIF
KGGGLRLISRLDCTNTWVQFSRLFIAGTPERSHILLOFVSFGSVTHLWMDFFY
FPQTTSLIFNLVYPYGAQTNATRFRLWQPYNNGKEEIIIDDFIDIGNNLNNPILL
LDTDFGPRDNWFFYPGGNIGLYCPSKGAPEDESAAMVFSNVEGHSITTRDLVS
NENTIIQFEINVCSTDSADPVRLEFRDFGATWHLLPLCYLHSSLSVLSICSTEH
HPSSTYACTTGGNRRVYHFGKHLKCGSVFRWYQGFYAGSQPYTWAIDNVYIGPQ
CEMVCYHGSCINGTKICDPTGSGTCKISTKNPDFLAKDFEGGLESDFLMSGGK
PBRKCGIILSSGNLNFEDGLRMLVTRDLDSHAREVQFMRLGCGKQVDPDRSQVL
LAYSLNGLSWLLQEFLEFSNNSNVGRYIALEMPKARSGSLRLRWQPSSENGFTSP
WYIDQTLIGNISGNTLBDDEFSLDSRKWLLHPGCTKMPVCGSTGDAIVFTKASTP
YVYVTTDIAVNESDFIQIDFAASCSVTDCYALEYSVDLGSWHPLRDCLPTNVEC
SYHLQRIILYSOTFNKWTITLPLPSYSTRSQATRFWRHOPAFDFKQOTWADNVYIGD
GCIDMCSHGRCVQSGCVCEQWGGGLYCDPEPSTLQTLKDNFNAPSNONWLTVSGG
KLSTVCGAVASGLAHFSGGCSRLVTVLDNLNTNAEFTIOFYMYGCLTTPSNRNOQVL
```

LEYSVNGITWNLMLLEIFVQYKPGFVNILLPPDAKEIATFRFWRQPRHGDLDQNDW
 AIDNVLISGADQRTWMLDTFSSAPVPQHERSPADAGPVGRIAPFEMFLEDKTISVNNW
 LFHDDCTVERFCSDPGVMLCGSHGDEHYAVTHDLTPTEWIMQKISYGCXVPEKI
 AQOIHVFTGDSVMSYLVPOCLPADPKCSGVSQPSVFETEGWKRTIYPLPEKI
 TGNPFRFYOKYSDVQWADNFYLGPGCLDNGHGDCLKEOCICDPGYSGPNCVLT
 HSLKTLKREFDSEELKPOLWMSLEGSTCEGVLAENTLYEGSTVQOATDOLD
 LRGAFLQYWGRTGSENNWTSCHRPVCKEGLVLDSTDGGITWTLHEMDFPKYISV
 RHDYILIPGALNTTKLRWQPFVINSGLVSVGERAQWALDNILIGGAENIPSOVLV
 DTFDDGSSHEENWSPYNAVRTAGFCNPSFHYWPKKKDKTKHNLSSRELIIQPG
 YMQFKTVSGEAKRTISQDLSHVMLEYTKDARSWSVOLVQCLPSSNSIGCSPPQFH
 EATYINAVNSSEKATIQLPDHSVSSATOFRIQKQETEKOSWADHWYIGEACPK
 LCSGHVCTTGAVCIDESPOGDDCSVESHELPSYIKDNFESARVTEANNETIOGGVI
 GSCGQLAPYAHGDSLYFNGCQIQRAATKPLDLITRASKINFLQIISPAQTDSCNSDL
 SGRHTVDKAVLLQYVNNGTIHWIAHQPKDFSTQORVSYNVPLEARMGVLLRWQ
 PRNGTGHQDQWALDHVEVLLVSTRKQNYMMNFROHLRHFYNNRRSLRRYP"
 10669. .11673
 11662. .11667

3'UTR
 polyA_signal
 BASE COUNT 2831 a 2985 c 2985 g 2872 t
 ORIGIN

Alignment Scores:
 Pred. No.: 6.56e-116 Length: 11673
 Score: 117.00 Matches: 117
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

US-09-897-438B-2 (1-117) x MMU24703 (1-11673)

Qy 1 GluGlnCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20
 |||||
 Db 970 GAGCAGTGTGGCACCACATCATGCGCAATGCTGTACCTTCTGTGAGCCGTACGGCCCT 1029
 Qy 21 ArgGluLeuThrThrCysLeuAsnThrThrThrAlaSerValLeuGlnPheSerIle 40
 |||||
 Db 1030 CGAGAGCTGACCACACATCGCTGACACAGACACAGCATCTGTCTCCATTTCCATT 1089
 Qy 41 GlySerGlySerCysArgPheSerTyrSerAspProSerIleThrValSerTyrAlaLys 60
 |||||
 Db 1090 GGGTCAGGATCATGTGCGATTTAGTTACTCTGACCCAGCATCTGTGTATACGCCAAG 1149
 Qy 61 AsnAsnThrAlaAspTrrPileGlnLeuCluLysIleArgAlaProSerAsnValSerThr 80
 |||||
 Db 1150 AACAAATCCGCTGATTGAGTTCAGTGGAGAAATATAGAGCCCTTCCAAATGTGAGCACA 1209
 Qy 81 ValIleHisIleLeuTyrLeuProGluAlaLysGlyGluSerValGlnPheGlnTrp 100
 |||||
 Db 1210 GTCATCCACATCTGTACTTCCCGGAGAACCCAAAGGGAGAGCGTGCAGTTCAGTGG 1269
 Qy 101 LysGlnAspSerLeuArgValGlyGluValTyrGluAlaCysTrrPalaLeu 117
 |||||
 Db 1270 AAACAGGACAGCCTCGAGTGGGTGAGTGTATGAGGCGTGTGCGCCCTG 1320

RESULT 3
 AC121878/c
 LOCUS AC121878 163985 bp DNA linear ROD 21-JUN-2002
 DEFINITION Mus musculus clone RP24-131C8, complete sequence.
 ACCESSION AC121878
 VERSION AC121878.1 GI:21039993
 KEYWORDS HTG.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 163985)
 McPherson, J.D. and Waterston, R.H.
 TITLE The sequence of Mus musculus clone
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 163985)
 McPherson, J.D. and Waterston, R.H.
 AUTHORS Direct Submission
 TITLE Submitted (21-MAY-2002) Genome Sequencing Center, 4444 Forest Park
 JOURNAL

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Parkway, St. Louis, MO 63108, USA
 3 (bases 1 to 163985)
 McPherson, J.D. and Waterston, R.H.
 Direct Submission
 Submitted (21-JUN-2002) Genome Sequencing Center, 4444 Forest Park
 Parkway, St. Louis, MO 63108, USA

----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 Contact: submissions@watson.wustl.edu
 ----- Project Information -----
 Center project name: M_BB0131C08

FEATURES
 Source
 1. 163985
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="RP24-131C8"
 BASE COUNT 50392 a 33960 c 32422 g 47211 t
 ORIGIN

Alignment Scores:
 Pred. No.: 4.1e-37 Length: 163985
 Score: 45.00 Matches: 45
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 38.46% Indels: 0
 DB: 10 Gaps: 0

US-09-897-438B-2 (1-117) x AC121878 (1-163985)

Qy 73 ArgAlaProSerAsnValSerThrValIleHisIleLeuTyrLeuProGluLualalys 92
 |||||
 Db 15382 AGAGCCCTTCCAAATGTGAGCAGCATCATCCATCTGTACCTCCCGAGGAGCCAAA 15323
 Qy 93 GlyGluSerValGlnPheGlnTrpLysGlnAspSerLeuArgValGlyGluValTyrGlu 112
 |||||
 Db 15322 GGGGAGAGCGTGCAGTTCAGTGGAAACAGGACAGCCCTGCGAGTGGGTGAGGTGTATGAG 15263
 Qy 113 AlaCysTrrPalaLeu 117
 |||||
 Db 15262 GCCTGCTGGGCCCTG 15248
 RESULT 4
 AC023062 183156 bp DNA linear HTG 05-MAY-2000
 LOCUS AC023062
 DEFINITION Mus musculus chromosome 5 clone RP23-466D2 strain C57BL6/J, WORKING
 DRAFT SEQUENCE, 8 unordered pieces.
 AC023062
 VERSION AC023062.2 GI:7709916
 KEYWORDS HTG; HTGS, PHASE1; HTGS_DRAFT.
 SOURCE Mus musculus.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 183156)
 Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W.,
 Bouffard, G.G., Dietrich, N.L., Eagle, W.O., Gupta, J., Ho, S.-L.,
 Huang, M.C., Idol, J., Lee-Lin, S.-Q., Maduro, O.L., Maduro, V.B.,
 Mastrian, S.D., McCloskey, J.C., Morse, E., Ojodu, M.A., Pearson, R.,
 Stantripop, S., Summers, T.J., Thomas, J.W., Thomas, P.J.,
 Tongson, E.E., Touchman, J.W., Tran, J.T., Vogt, J.L., Walker, M.A.,
 Wetherby, K.D., and Green, E.D.
 TITLE NISC Mouse Sequencing Initiative
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 183156)
 Green, E.D.
 TITLE Direct Submission
 JOURNAL Submitted (08-FEB-2000) NIH Intramural Sequencing Center, 8717
 Government Circle, Gaithersburg, MD 20877, USA
 On May 5, 2000 this sequence version replaced gi:6939141.

----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.mouse@nih.gov
----- Project Information
Center project name: Y1
Center clone name: 466D02

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 175371 bases at least Q40
Consensus quality: 178474 bases at least Q30
Consensus quality: 180133 bases at least Q20
Insert size: 179000; agarose-fp
Insert size: 188000; pulse-field-gel
Insert size: 183156; sum-of-contigs
Quality coverage: 6.96x in Q20 bases; agarose-fp
Quality coverage: 6.63x in Q20 bases; pulse-field-gel
Quality coverage: 6.80x in Q20 bases; sum-of-contigs

*** NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence, record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```
1 2374: contig of 2374 bp in length
   gap of unknown length
* 2375 4490: contig of 2116 bp in length
   gap of unknown length
* 4491 14959: contig of 10469 bp in length
   gap of unknown length
* 14960 31072: contig of 16113 bp in length
   gap of unknown length
* 31073 50119: contig of 19047 bp in length
   gap of unknown length
* 50120 78196: contig of 28077 bp in length
   gap of unknown length
* 78197 109874: contig of 31678 bp in length
   gap of unknown length
* 109875 183156: contig of 73282 bp in length.
```

FEATURES
source
1. 183156
/organism="Mus musculus"
/strain="C57BL6/J"
/db_xref="taxon:10090"
/chromosome="5"
/clone="RP23-466D2"
/clone_lib="Rpci mouse BAC library 23"
BASE COUNT 53719 a 37080 c 37302 g 55052 t 3 others
ORIGIN

Alignment Scores:
Pred. No.: 4,52e-37 Length: 183156
Score: 45.00 Matches: 45
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 38.46% Indels: 0
DB: 2 Gaps: 0

US-09-897-438B-2 (1-117) x AC023062 (1-183156)

Qy 73 ArglaProSerAnValSerThrValIleHisIleLeuTyrLeuProGluAlaLys 92
|||||
Db 173623 AGACCCCTTCCAAATGTGAGCACAGTACATCATCTGTAACCTCCCGAGGAGCCAAA 173682
|||||
Qy 93 GlyGluSerValGlnPheGlnTrpLysGlnAspSerLeuArgValGlyGluValTyrGlu 112
|||||
Db 173683 GGGGAGAGCGTGCACTTCCAGTGAAGAAACAGACAGCCTGCGAGTGGGTGAGGTGTATGAG 173742
|||||

Qy 113 AlaCystTrpAlaLeu 117
|||||
Db 173743 GCCTGCTGGGCCCTG 173757
RESULT 5
AB049473
LOCUS
DEFINITION Rattus norvegicus mRNA for reelin, complete cds.
ACCESSION AB049473
VERSION AB049473.1 GI:17221617
KEYWORDS
SOURCE Rattus norvegicus juvenile cerebellum cDNA to mRNA.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 Kikkawa,S. and Terashima,T.
TITLE rat reelin (Reln) complete CDS
JOURNAL Published Only in Database (2001)
REFERENCE
2 (bases 1 to 11187)
Kikkawa,S. and Terashima,T.
TITLE Direct Submission
JOURNAL Submitted (29-SEP-2000) Satoshi Kikkawa, Kobe University School of
Medicine, Department of Anatomy: Chuo-ku, Kusunoki-cho 7-5-1, Kobe,
Hyogo 650-0017, Japan (E-mail:skikkawa@med.kobe-u.ac.jp,
Tel:81-78-382-5325, Fax:81-78-382-5328)
FEATURES
source
1. 11187
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/tissue_type="cerebellum"
/dev_stage="juvenile"
1. 11187
/gene="Reln"
/gene="Reln"
/codon_start=1
/product="reelin"
/protein_id="BAB78470.1"
/db_xref="GI:17221618"
/translation="MERGCAWPTLVAVLLLLATLRRAATGYPRFSPFFLCTH
HGELEGDGEQEVLSIHAGNPTLVVPGQYHVTISTFFDLGLVTLGYTSTSTQS
SQSIGSSAGFGFIMSDHOFQCMFVSVASHVSHLPTTNLSPVLTAPAGTSCVNF
ATATHRGQVLEKDALAQCEOGATEATAYSHLAEIHSDSVILRDDFSYHOLELN
NIWAECNCDTGEQCGTIIHGNVATCEPYGPRELITTTTNTTASVLFQSGSCSR
FSYSDPSIIYSYAKNNTADWIKLEIRAPSNVSTIIHLYLPEDAKGNVQFQWKDS
LHVGEVYACWALDNILVINSARQVILESDLPDVTGNMFPFGATVKHSCQSDNA
IYFPHGEGQLNFATTTRDVLSTEDIQEQWSEFEFESQPTGMDTLGAVGSECTIEG
LSLVFLKDGKRLCTPYMDTGYGNLFYFAMGTCPCDGDHENDVLYAKIEGKEH
IALDTLSYSYKVTPLVSVVNPLOTPATKCLCKROKHQGNQYVAVDFHVLPTL
PTMSHWIQFSLNLCGTHQPCNSVLESTNHNRSWSLLHTECLPEICAGPHLPST
IYSSENISGNRVITFPLNALTDRIRMRQTGPILGNMWAIDNVIIGPSCULKFCG
RGQCTHRGCKDPPGSPACEMASQTFPMFISFSGSRSLSSYHNFYISRGAEVFGC
GVLASGLKFNKDRRLITSLFSSQSRFLQTLRLGSKSVLSTCRAPDQEGVL
LHYSDNGITTKLLEHYSYLVNHEPTIISVELPDDAQFGIOFRWQPHSSQGEDVW
AIDEILMTSVLFNSISLDTNLVENVQSLGFLGNIQPCYGHDMTLCFTGDSKLASS
RYVEITCSMOIGASYMTQESLVMGCGQKYPHMDNVKLEYSTNHGLTWHLVQDECLPS
MPSQCFTSASIHASEFTQWRVITLVKQKTSWGAIRFWSQSYTTAQDEWALDDIY
IGQCCPNMCSGDHGCRCQDQGYGTECHPEAALPTIMSDFNENPSSWSDWDQEV
IGGEVVKPEQCGVSSGSLYFSKAGKQLSWDLDTSWDFVQFYIIGGSAACN
KPDSEEGVLLQYSSNGGIQWHLAEWYFSDFKPRFVLELPAATAKTCTCFRFRWQP
VFGEDYDQWAVDDIIILSEKOKOVIIPVNPVLTPONFEKPAFDYPIINOMSVMLAN
EQMAKNDSCATTPSAMVFGKSDGDFRFAVTRDLTLKPGVLOFLKLNIGCASQFSSATP
VLLOYSHDAGMSFLVKEGCFPASAGKCGEGNSRELSEPTVITGTGFEWRVITAIIP
RLASAKTRFRWIESSQKNVPFGLDGYIISEPCSYSCSGHGDICISGFCFDLGYT
AAQGTCSNIPNHEMFDEGKPLWYKISGQVCTGCCTGSLDGRSLYFNGLKRE
ARTVPLDTRNIRLQVYIQTGSKTSGITCKPRARNEGLVVOYSDNGLIWLHLLRLD
FLSFLPQIISIDLPREKATPATFRWQPHQCKHSAOWALDDVLIGVNDSSQTFQD
KPDGSLDLQANRYRLOGGOVDLIDLSMDTALLFTENIGKPRYAEWDFHVSASFQF
DKSMGSKFPFSAHVSQLOYSLNNGKDMHPVTEECVPPTIGCVHYTSSYTSERFQN
WRVTVYLPATNSPRTFRFRWQITMTGADAWAIDNVLASGCPWLCSGRCGICDSGR

CYCDGFGGPFPCVVPVPLPSLLKDDFNGNLHPDLWPVYGAERGNLNGETIKSGTSLI
FKGEGRLMFLCDTNTMYVQFSLRFLAKGTPERSKSLIQSSINGGVYWRKLMDEF
YFQPTTLLIFINVPDYSIAQTNATRLRQWPNYNGKKEIWIIDDIIDCDNLNNPWF
LIDTFDGFREDFNMFYPGNGTGLYCPYSKGAPEDSAMVFSVNEIGHSTTRDLS
VNEHTIOFIEINVCGSTYDSSADVRLEFSDFGATWHLILLPLCYHSSSLVSLCSTE
HHPSSSTYAGTTQWREYVHFCKLHLCGSVRFRWQYGFVPAGSQPVTAIDNVYIGP
QCEMCCGHCSCVNGKICDPCYSGPTCKISTKPNDFLKDQFEGOLESDRELLMSGG
KPSKCGIILSSGNLFFNEDGLRMLVTRDLDSHARFVQFFMRGLGCGKGVDPDRSPV
LQYISNGLSLLEQFIFSNSSNGRYIALEMLPKARSGSTRLRWQFSENGHFIYS
PWYDILGGINISNTVLEDFSTLDSRKLWLLHPGGTRKMPVCGSDALVFTEKAST
RYVWTDIAVNSDFQIDFAASCSTVDSCYALEEYSVDLSWHLPLVRDCLPTNVE
CSRYHLQRIIVSDTEKNWTRITLPLPAYTRSOATRRHQAPEDKQOOWAIDNVYIG
DECLDWCSGHRCIQSCYCDGEGWGLYCDDEPSTLPTQLKNFNAPSNQWLTNG
GKLSYVCGAVASGLAHFSGGSKLLVTDNLNNAEFTQFYMYGCLLTPSRNNGV
LLEYSYNGITLWLLIEFYDYQSKFGEVNIILPPDAKEIGTRFRMRQRHGLDQND
WAIDNLIIGSADQRTVMDLTFSSAPVQHERSPADAGPVGFIAFDMFMDKTSVNE
WVPHDCTYERFDCSDPGVMLCSDHGRVYAVTHDITPTENWIMGFKISVCGKVPK
IAQNOITHOFTSTDFGWSVLYPQCLPADPKSGTYSOPSVPFPTKGRKRTIYPLPES
LGNVPYFRFYOKYSDOVAIDNFIYLGPCGLDNCGGHGDCLKEQCICDPYSGPHCYL
THLTKFLERFDSBEIKFDLMSLEGGSTCTECGIAENTALFYGGSTVRAITQDL
DLRGAFLOYWGRIGSENNTSCHRPVCRKEGVLLDYSKDGGITWLLHEMDFOKYIS
VRHDYTLIPPEGALTNTLRWQPPFVLSNGLVSGVERAQWALDNLILGAEINLSQ
VDFDDEGSHENWFYNAVRTAGFCGNPSPHLYPNKKDKTTHALSSRLIITOP
GYMWOPIVVGCEATSCGDLHSVMLEYTKDARSDSQVLQVOTCLPSSNSIGCSPOF
HEATYINAVNSSWKRITITQLPDRHVSSTATQFRWQKGETEKQSWAIDHYVIGERAC
RLCSGHGCTTGAVCICDESFGDDCSVFSHELPSYIKONFESARVTEANWETIQGGA
IGSGGOLAPYAGHDSLYNGCOIQRAQKPLDLTRASKIMEVLQIGSTAQDSCNSD
LSGPHTVDAVLLVSYNHTIWHVIAHQPKDFTQAQRYSYNVPLEARMKGVLLRW
QPRHNGTGHQWALDHYEVVVLVSTRKQNMNNEFRQHLRHFYNNRRRSURRYP"

BASE COUNT 2645 a 3029 c 2930 g 2583 t

ORIGIN

Alignment Scores:
Pred. No.: 1.08e-18 Length: 11187
Score: 27.00 Matches: 27
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 23.08% Indels: 0
DB: 10 Gaps: 0

US-09-897-438b-2 (1-117) x AB049473 (1-11187)

Qy 28 LeuAsnThrThrAlaSerValLeuGlnPheSerIleGlySerGlySerCysArgPhe 47
|||||
Db 1112 CTGAACACGACACAGCATCGCTCCAGTTTCCATTGGTCAGGATCCTGTCGATT 1171
Qy 48 SerTyrSerAspProSerIle 54
|||||
Db 1172 AGTACTCTGACCCACGATC 1192.

RESULT 6
AX410790 11580 bp DNA linear PAT 14-JUN-2002
LOCUS
DEFINITION Sequence 3437 from Patent WO0229103.
ACCESSION AX410790
VERSION AX410790.1 GI:21443495
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Alvares.C., Horne.D., Peres-da-Silva,S. and Vockley,J.G.
Gene expression profiles in liver cancer
Patent: WO 0229103-A 3437 11-APR-2002;
GENE LOGIC INC (US)
FEATURES
Location/Qualifiers
source
1. .11580
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="EMBL/GenBank Accession No. U79716"
BASE COUNT 3014 a 2696 c 2753 g 3116 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 1.12e-18 Length: 11580
Score: 27.00 Matches: 27
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 23.08% Indels: 0
DB: 6 Gaps: 0

US-09-897-438b-2 (1-117) x AX410790 (1-11580)

Qy 28 LeuAsnThrThrAlaSerValLeuGlnPheSerIleGlySerGlySerCysArgPhe 47
|||||
Db 941 CTTAATACACACAGCTTCTGCTCCAAATTTCCATTGGTCAGGTTTCATGTCGCTTT 1000
Qy 48 SerTyrSerAspProSerIle 54
|||||
Db 1001 AGTATTACACCCACGATC 1021

RESULT 7
HSU79716 11580 bp mRNA linear PRI 25-FEB-1997
LOCUS
DEFINITION Human reelin (RELN) mRNA, complete cds.
ACCESSION U79716
VERSION U79716.1 GI:1743884
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 11580)
DeSilva,U., D'Arcangelo,G., Braden,V.V., Chen,J., Miao,G.G.,
Curran,T. and Green,E.D.
The human reelin gene: isolation, sequencing, and mapping on
chromosome 7
JOURNAL Genome Res. 7 (2), 157-164 (1997)
MEDLINE 97202106
PUBMED 9049633
REFERENCE 2 (bases 1 to 11580)
DeSilva,U., D'Arcangelo,G., Braden,V.V., Chen,J., Miao,G.G.,
Curran,T. and Green,E.D.
Direct Submission
JOURNAL Submitted (26-NOV-1996) National Center for Human Genome Research,
National Institutes of Health, 49 Convent Drive, MSC4431, Bethesda,
MD 20892, USA
FEATURES
Location/Qualifiers
source
1. .11580
/organism="Homo sapiens"
/db_xref="taxon:9606"
1. .11580
/chromosome="7"
gene
/gene="RELN"
176..10558
/gene="RELN"
/codon_start=1
/product="reelin"
/protein_id="AAC51105.1"
/db_xref="GI:1743885"
/translation="MERSGAROTFELLALLGATLRLARAAAGYVPRFSPFFLCTHHG
ELSGDEQGEVLISHIAGNPTYYVPGQEVHVITSTFEDGLLVTLCTSTSVQASQ
SICGSAFEGFIMSDHQFGNQFMCSSVAVSHVSLPTNLNLSFIWIAPAGTCVNFNAT
ATHRGVIFKDALAQQLCEGAPTDVTHPLAEITHSDSILKDDFDSYHQDLQNPNI
WVECNQCTGECGCAIMHNAVTFCEPYGPRELITTLGLNTTASVLFQSIGSGCRFS
YSDPSITVLYAKNSADWIOLEKIRAPSNVSTIHLILPEDAKGVQFQWQENLR
VGEVYBACWALDNLILNSAHROVLEDSLDVDTGNLWLPFCATVKKSCQSDNSIY
FHNGESSENFATRDVLDSTEDIQWSEFEFESQPTGHDVGLAVLGTGCGTIESLS
MVFLDGGRKKLCTPSMDITGYNLRFYVMGGICDPGNSHENDVIAFKIEGKEHIT
LDTLSYSSKVPVLSVWINPELQTPATKFLQKQKHQHNRMVAVDFHVLPLVPS
TMSHMIQFISINLGNATHQNGSVLEFSTNHRGSWSLLHTEPELCAAGPHLPHSTVY
SSENYSQWNRITITPLNNAALTRNTRTRWOTGPIILNMWALDNVYICPCLKPCSGRG
QCTRHCCKDPPGSGPACEMASOTFMFISESGSSRLSSYHNFTYIRCAEVSFGGV
LASGKALVFNKEGRRLITISFLDSSSRFLOFTLRGLGSKSVLSTCRAPQPGEGVLLH
YSYDNGITWKLLEHYSTLSYHPEPRISISVELPQDAKQFGIQFRWQYHSSQREDVWAI

DEIMTSVFNLSISLDFTNLVEVTQSLGFLYGNVQVPCGHWDWLCFTGDSKLASSMRY
VFQSMOIGASVYQISLVWCGCKQVTPHMDNOVKLEYSNTNHLGTLWHLVOECLPSMP
SCOEFTSASVYHASEETQWRVVLVLPQKTSWSSATRRWSQVYTAQDEWALDSIYIG
QCPNMCSCGSDHICRCDOQYQGTCEHPEALSTINSDFENONGWESDQWEVIG
GEIYBPQGGCVIISGSSGLYFSKAGKQLVSWDLSTWVDFVQFYTOIGESASCKRP
DSREGLVLLQYNNNGTOMHLLAEYFSDSKPRFVYLELPAKAACTPCTFRFRWQVFE
SGEDYDCAWDDIILISEKQKIPIVNTPLQNFYKPAFYPMQMSVWMLANEG
MVNETFCAMTFSAMIFGKSDGRFAVTRDLTLKPGYVLOFKLNICANOFSTAPVL
LOYSHDAGSMFLVKGCYFASAGKCEGNSRELSPTMYHTGDFEETWITIVIPRS
LASKTRFWIQESSOKNVPFGLGVYISEPCPSYCSGCDICSGVCGDLYTAA
QGTCSVNPNNHNEFRFEGKLSPLWKITGAQVGGCGTLNDKCSGLYFNGPKREAR
TVPELITNIRLRYFYTOIGKTSGITCIKPRTRNEGLIYOVSNDNGIMLHLRELDFM
SFLPDIITSLDPODQATPATAFPMQPOHQKHSQAQWALDDVILGNDSQTFQDFK
DSGEDYDCAWDDIILISEKQKIPIVNTPLQNFYKPAFYPMQMSVWMLANEG
SMGSKPFSNSHVSQVLYSLNGLKMDHLTEECVPTTIGLHVTESSITVSERFQNK
RTVYLPSTISPTFRFTQANTVGAOSWADNVVLASGCPMCGSGICDAGCV
CDRGFGPCVCPVPLPSILKDFENGLHPLDLPVENYGAERGNLNETIKSGTSLIFK
GEGRLMISDLDOCTNMTWVFSRFLIAKSTPERSHILQFISGGITWHLMDSEFYF
PQTNILFINVPLPYTAQTNATFRWQYNNGKKEIWIWDDFIIDGNVNNPVML
DTDFGPREDNMFYFGNIGLYCPTSSKGADEDSAMVFSVEGHSITTRDLNVN
ENTIIQFEINVGCTSSADPVRLEFSRDFGATWHLPLCYHSSSHVSLCSTEH
PSSTYAGTMQWRREVHFKLHLCGSVFRWYQGFYFAGSQPVTWADNVYIGPQC
EEMNGGSCINGTKICDPYSGPTCKISTKNPDELKDDFEGQLSDREFLMSGKGP
SRKGLSSGNLFFNEDGLRMLTRDLSHARFVOFEMRLCGCGKGVDPDRSQPVLL
QYSLNGLSWLLQEEFLFNSSNVGRYIALEIPLKARSGSTRLMWQPSENGHFYSPW
VIDQILIGNISGNTVLEDDFTLDRSKWLLHPGKMPVCGSTGALVFIKASTRY
VSTDVAVNEDSLQIDFAASCSTDCSVAIEYSDVLGSLWHLVRLDCLPTNVECS
RYHLRLVSDTFENKTRITVLPPLPYTRQATFRWHOPAPEDKQOTWALDNVYIGDG
CIDMCSGRCICGNCVCDQOMGLYCDDEPESLPTOLKDNENRAPSNNWLVNCGK
LSTVCAVASGMALHFGSGCSRLVTVDLNLNAEFIQFIYFGLITPNNRQGVLL
EYVNGITWNLMEIFYDQYSKPGEVNIPLPDPAKEIATFRFRWQPRHGDLDNDWA
IDNVLLSGAQRTVMTDFTSSAPYHESPADAGPVGRIAFDMEDKTSYNEHL
FHDCTYERFCDSPDGMCLGSHDREVAVTHDLTPTCEWTMOFKISVCGKVKEXIA
QNOIHVOYSTDFGVSNVLYPOCLPADKCSGSVSPSVFFPTKWKRTYPLPESLV
GNPVREFYQKYSMDQWADNFIYLPGLDNCRGHCLREQCICDPYSGPNCYLTH
TLKTFLEKRFDESEIKRDLWMSLEGSTCEGILAEEDTALFGLGTVRQAVTQDLDL
RGAKFVLEKRGSENNMTSCHRIKCBEGVLLDYSTDGGITWTLHEMDYQKYSVR
HDYILLPEALNTLRLWMPQFVISGVSVGERAOWALDNILIGGAEINPLQVLD
TFDEGTSHEENMSFYPAVRTAGFCGNPSFHLWPNKKDKKTHNALSSRELIIQPGY
MMQFKIVGCEATSCGDLHSLVMEYTKDARSWSLVQTOCLPSSNSIGCSPFPKE
ATLYNSVSSSKRITVLPDPAHVSATOFRTWKGEETEKOAWIDHVVIGEACPL
CSGHGYCTTGATCDESPQDDCSVESDLPYIKDNFESARVTEANNETQGGVIG
SGCGQLAPYAHGDSLFENGCQIQAAATKEDLDRASKIMFVIOIGMSQTDSCNSDLS
GPHAVDKVLLQYSVNGGITHVIAHQPKDFQAOQVSYNVPLEARMKGLLRWMP
RHNGTGDQWALDHVLEVLVSTRKQNMNFSRQGLRHFYNNRRSLRRYP"

BASE COUNT 3014 a 2696 g 3116 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 1.12e-18 Length: 11580
Score: 27.00 Matches: 27
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 23.08% Indels: 0
DB: 9 Gaps: 0

US-09-897-438b-2 (1-117) x HSU79716 (1-11580)

QY 28 LeuAsnThrThrAlaSerValLeuGlnPheSerIleGlySerGlySerCysArgPhe 47
|||||
Db 941 CTTAATACACACAGAGCTTCTGCTCCCAATTTCCATTGGGTCAGGTTTCATGTCGCTTT 1000
|||||
QY 48 SerTyrSerAspProSerIle 54
|||||
Db 1001 AGTTATTACAGACCCAGCATC 1021
|||||

RESULT 8
G30936
LOCUS
DEFINITION
G30936 Eric D. Green Homo sapiens STS genomic, sequence tagged
site
ACCESSION
G30936
VERSION
G30936.1 GI:1923209

KEYWORDS
SOURCE
ORGANISM

STS.

Homo sapiens.

Homo sapiens
Homo sapiensEukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 11580)

AUTHORS

Bouffard,G.G., Iyer,L.M., Idol,J.R., Braden,V.V., Cunningham,A.F.,
Weintraub,L.A., Mohr-Tidwell,R.M., Peluso,D.C., Fulton,R.S.,
Leckie,M.P. and Green,E.D.A collection of 1614 human chromosome 7-specific STSs
Genome Res. 7 (1), 59-64 (1997)

97189344

9037602

REFERENCE

2 (bases 1 to 11580)

Green, E.D.

Human chromosome 7 STSs (1997)

Unpublished (1997)

On Apr 3, 1997 this sequence version replaced gi:1706935.

Synonyms: RELN

GDB.DSEG: RELN

Contact: Eric D. Green

Genome Technology Branch

National Human Genome Research Institute/NIH

49 Convent Dr., MSC4431, Bldg. 49, Rm. 2A08, Bethesda, MD 20892

Tel: 3014020201

Fax: 3014024735

Email: egreen@nhgri.nih.gov

Primer A: CTCCTACCTTCTCTGAGGACGCCAA

Primer B: AAGCAGTTGCTGTGTCCACTG

STS size: 183

PCR Profile:

Presoak: 0 degrees C for 0.00 minute(s)

Denaturation: 92 degrees C for 0.17 minute(s)

Annealing: 55 degrees C for 1.00 minute(s)

Polymerization: 72 degrees C for 1.00 minute(s)

PCR Cycles: 35

Thermal Cycler: PerkinElmer 9600

Protocol:

Template: 30-100 ng

Primer: each 1 uM

dNTPs: each 200 uM

Taq Polymerase: 0.05 units/ul

Total Vol: 10 ul

Buffer:

MgCl2: 1.5 mM

KCl: 100 mM

Tris-HCl: 10 mM

NH4Cl: 5 mM

pH: 8.6

This STS was developed from sequence determined by another investigator. See GenBank record: U79716 For additional information about the NHGRI chromosome 7 mapping project, see <http://www.nhgri.nih.gov/DIR/STB/CHR7>. Also see Genomics 11:548-64 (1991) [MUID=92128937].

FEATURES

Source

1. .11580
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="7"
/clone_lib="Eric D. Green"

gene

1. .11580
/gene="RELN"
1112. .1294
/gene="RELN"
1112. .1134
/gene="RELN"

primer_bind

3014 a 2696 c 2753 g 3116 t
complement(1272..1294)

primer_bind

BASE COUNT 3014 a 2696 c 2753 g 3116 t
ORIGIN

Alignment Scores:

Pred. No.: 1.12e-18 Length: 11580
 Score: 27.00 Matches: 27
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 23.08% Indels: 0
 DB: 11 Gaps: 0

US-09-897-438B-2 (1-117) x G30936 (1-11580)

Qy 28 LeuAsnThrThrAlaSerValLeuGlnPheSerIleGlySerGlySerCysArgPhe 47
 |||||
 Db 941 CTTAATACACACAGCTCTGTCTCCAAATTTCCATTTGGTCAGGTTTCATGTCGCTTT 1000

Qy 48 SerTyrSerAspProSerIle 54
 |||||
 Db 1001 AGTTATTAGACCCAGCATC 1021

RESULT 9
 G30938 11580 bp DNA linear STS 28-SEP-1998
 LOCUS SWSS3176 Eric D. Green Homo sapiens STS genomic, sequence tagged
 DEFINITION

ACCESSION G30938
 VERSION G30938.1 GI:1923211
 KEYWORDS STS.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Bouffard, G.G., Iyer, L.M., Idol, J.R., Braden, V.V., Cunningham, A.F., Weintraub, L.A., Mohr-Tidwell, R.M., Peluso, D.C., Fullon, R.S., Leckie, M.P. and Green, E.D.
 TITLE A collection of 1814 human chromosome 7-specific STSs
 JOURNAL Genome Res. 7 (1), 59-64 (1997)
 MEDLINE 97189344
 PUBMED 9037602

REFERENCE
 AUTHORS 2 (bases 1 to 11580)
 Green, E.D.
 TITLE Human chromosome 7 STSs (1997)
 JOURNAL Unpublished (1997)
 COMMENT On Apr 3, 1997 this sequence version replaced gi:1706937.
 Synonyms: RELN
 GDB_DSEG: RELN
 Contact: Eric D. Green
 Genome Technology Branch
 National Human Genome Research Institute/NIH
 49 Convent Dr., MSC4431, Bldg. 49, Rm. 2A08, Bethesda, MD 20892
 Tel: 3014020201
 Fax: 3014024735
 Email: egreen@nhgri.nih.gov
 Primer A: TGTGGCTGATTACTAACAG
 Primer B: TGGGCTTTTATAGCTTG
 STS size: 238
 PCR Profile:

Presoak: 0 degrees C for 0.00 minute(s)
 Denaturation: 92 degrees C for 0.17 minute(s)
 Annealing: 50 degrees C for 1.00 minute(s)
 Polymerization: 72 degrees C for 1.00 minute(s)
 PCR Cycles: 35
 Thermal Cycler: PerkinElmer 9600

Protocol:
 Template: 30-100 ng
 Primer: each 1 uM
 dNTPs: each 200 uM
 Taq Polymerase: 0.05 units/ul
 Total Vol: 10 ul

Buffer:
 MgCl2: 1.5 mM
 KCl: 100 mM
 Tris-HCl: 10 mM
 NH4Cl: 5 mM

pH: 8.6

This STS was developed from sequence determined by another investigator. See GenBank record: U9716 For additional information about the NHGRI chromosome 7 mapping project, see <http://www.nhgri.nih.gov/DIR/CTB/CHR7>. Also see Genomics 11:548-64 (1991) [MUID=92128937].

FEATURES
 Location/Qualifiers
 source
 1..11580
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="7"
 /clone_lib="Eric D. Green"
 1..11580
 /gene="RELN"
 11314..11551
 /gene="RELN"
 11314..11333
 /gene="RELN"
 primer_bind
 complement(11535..11551)
 primer_bind 3014 a 2696 c 2753 g 3116 t 1 others
 BASE COUNT
 ORIGIN

Alignment Scores:
 Pred. No.: 1.12e-18 Length: 11580
 Score: 27.00 Matches: 27
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 23.08% Indels: 0
 DB: 11 Gaps: 0

US-09-897-438B-2 (1-117) x G30938 (1-11580)

Qy 28 LeuAsnThrThrAlaSerValLeuGlnPheSerIleGlySerGlySerCysArgPhe 47
 |||||
 Db 941 CTTAATACACACAGCTCTGTCTCCAAATTTCCATTTGGTCAGGTTTCATGTCGCTTT 1000

Qy 48 SerTyrSerAspProSerIle 54
 |||||
 Db 1001 AGTTATTAGACCCAGCATC 1021

RESULT 10
 AC124933/c
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-466N17, *** SEQUENCING IN PROGRESS
 AC124933
 ACCESSION AC124933.2 GI:21952716
 VERSION
 KEYWORDS HTG; HTGS_PHASE1.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE
 1 (bases 1 to 183641)
 Muzyly, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C., Alsbrooks, S.L., Amaratunga, H.C., Are, J.R., Ayele, M., Banks, T., Barberia, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denny, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hoques, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,

```

Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karisson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louissege,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenwo,S., Oguh,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.I., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,A.,
Sodergren,E., Sonaika,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Woodson,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
2 (bases 1 to 183641)
Worley,K.C.
Direct Submission
Submitted (20-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 183641)
Worley,K.C.
Direct Submission
Submitted (31-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 24, 2002 this sequence version replaced gi:21490053.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KAVI
Center clone name: CH230-466N17
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 111724 bases at least Q40
Consensus quality: 120970 bases at least Q30
Consensus quality: 128458 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 88 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1108 1207: contig of 1107 bp in length
1208 2460: contig of 1253 bp in length
2461 2560: gap of unknown length
2561 4004: contig of 1444 bp in length
4005 4105: gap of unknown length
4105 5382: contig of 1278 bp in length
5383 5483: gap of unknown length
5483 6874: contig of 1392 bp in length
6875 6974: gap of unknown length
6974: gap of unknown length
8071: contig of 1097 bp in length
8171: gap of unknown length
8171: contig of 1307 bp in length
9478: gap of unknown length
9578: gap of unknown length
11061: contig of 1483 bp in length
11161: gap of unknown length
12742: contig of 1581 bp in length
12842: gap of unknown length
14504: contig of 1662 bp in length
14604: gap of unknown length
15734: contig of 1130 bp in length
15834: gap of unknown length
16958: contig of 1124 bp in length
17058: gap of unknown length
18704: contig of 1646 bp in length
18804: gap of unknown length
20420: contig of 1616 bp in length
20520: gap of unknown length
21648: contig of 1128 bp in length
21748: gap of unknown length
23177: contig of 1429 bp in length
23277: gap of unknown length
24760: contig of 1483 bp in length
24860: gap of unknown length
26759: contig of 1899 bp in length
26859: gap of unknown length
28299: contig of 1440 bp in length
28399: gap of unknown length
29592: contig of 1193 bp in length
29692: gap of unknown length
30870: contig of 1178 bp in length
30970: gap of unknown length
32642: contig of 1672 bp in length
32742: gap of unknown length
34386: contig of 1644 bp in length
34486: gap of unknown length
36084: contig of 1578 bp in length
36184: gap of unknown length
37398: contig of 1234 bp in length
37498: gap of unknown length
39608: contig of 2110 bp in length
39708: gap of unknown length
41480: contig of 1772 bp in length
41580: gap of unknown length
42759: contig of 1179 bp in length
42859: gap of unknown length
44711: contig of 1852 bp in length
44811: gap of unknown length
46208: contig of 1397 bp in length
46308: gap of unknown length
47767: contig of 1459 bp in length
47867: gap of unknown length
49227: contig of 1360 bp in length
49327: gap of unknown length
50797: contig of 1470 bp in length
50937: gap of unknown length
51928: contig of 1031 bp in length
52028: gap of unknown length
53105: contig of 1077 bp in length
53205: gap of unknown length
54546: contig of 1341 bp in length
54646: gap of unknown length
55820: contig of 1174 bp in length
55920: gap of unknown length
57635: contig of 1715 bp in length
57735: gap of unknown length
58942: contig of 1207 bp in length
59042: gap of unknown length
60341: contig of 1299 bp in length
60441: gap of unknown length
61693: contig of 1252 bp in length
61793: gap of unknown length
63315: contig of 1522 bp in length

```

```
* 63316 63415: gap of unknown length
* 63416 64571: contig of 1156 bp in length
* 64572 64671: gap of unknown length
* 64672 66083: contig of 1412 bp in length
* 66084 66183: gap of unknown length
* 66184 67516: contig of 1333 bp in length
* 67517 67616: gap of unknown length
* 67617 69820: contig of 2204 bp in length
* 69821 69920: gap of unknown length
* 69921 71262: contig of 1341 bp in length
* 71263 71361: gap of unknown length
* 71362 72775: contig of 1414 bp in length
* 72776 72875: gap of unknown length
* 72876 73298: contig of 2423 bp in length
* 73299 73398: gap of unknown length
* 73399 77094: contig of 1696 bp in length
* 77095 77194: gap of unknown length
* 77195 78834: contig of 1640 bp in length
* 78835 78935: gap of unknown length
* 78936 79889: contig of 1055 bp in length
* 79890 80089: gap of unknown length
* 80090 81774: contig of 1685 bp in length
* 81775 81874: gap of unknown length
* 81875 83255: contig of 1381 bp in length
```

Alignment Scores:

```
Pred. No.: 2,66e-13 Length: 183641
Score: 23.00 Matches: 23
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 19.66% Indels: 0
DB: 2 Gaps: 0
```

US-09-897-438b-2 (1-117) x AC124933 (1-183641)

```
Qy 1 GluInCysGlyThrIleMetHisGlyAsnAlaValThrPheCysGluProTyrGlyPro 20
|||||
Db 92916 GAACAGTGTGCACCATCATGCATGGCAACGCTGTACCTTCTGTGACCATATAGTGCCT 92857
|||||
```

Qy 21 ArgGluLeu 23

|||||

Db 92856 CGAGAGTTG 92848

RESULT 11

AC095877

DEFINITION Rattus norvegicus clone CH230-10G13, *** SEQUENCING IN PROGRESS HTG 11-JUL-2002

***, 59 unordered pieces.

AC095877

AC095877.4 GI:21722955

KEYWORDS HTG; HTGS_Phrase1.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 126130)

Muzy, D.M., Adams, C., Adlo-Oduola, B., Ali-Osman, F.R., Allen, C., Alsbrooks, S.L., Amarantunga, H.C., Are, J.R., Ayele, M., Banks, T., Barbarella, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,

Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseg, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I., Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wlezyk, R., Woodson, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstock, G., and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 126130)
Worley, K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 126130)
Worley, K.C.
Direct Submission
Submitted (11-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jul 10, 2002 this sequence version replaced gi:20975941.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- project Information
Center project name: GDRS
Center clone name: CH230-10G13
----- Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye
Assembly program: Phrap; version 0.990329
Consensus quality: 70934 bases at least Q40
Consensus quality: 75985 bases at least Q30
Consensus quality: 80731 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 59 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
* 1 1342: contig of 1382 bp in length
* 1383 1482: gap of unknown length
* 1483 3046: contig of 1564 bp in length
* 3047 3146: gap of unknown length
* 3147 4489: contig of 1343 bp in length
* 4490 4589: gap of unknown length
* 4590 5786: contig of 1197 bp in length
* 5787 5886: gap of unknown length

5887 7115: contig of 1229 bp in length
 7116: gap of unknown length
 7216 8771: contig of 1556 bp in length
 8772 8871: gap of unknown length
 8873 10260: contig of 1389 bp in length
 10261 10360: gap of unknown length
 10361 11370: contig of 1010 bp in length
 11371 11470: gap of unknown length
 11471 12710: contig of 1240 bp in length
 12711 12810: gap of unknown length
 12811 14180: contig of 1370 bp in length
 14181 14280: gap of unknown length
 14281 16048: contig of 1768 bp in length
 16049 16148: gap of unknown length
 16149 17569: contig of 1421 bp in length
 17570 17669: gap of unknown length
 17670 18820: contig of 1151 bp in length
 18821 18920: gap of unknown length
 18921 20209: contig of 1289 bp in length
 20210 20309: gap of unknown length
 20310 21633: contig of 1324 bp in length
 21634 21733: gap of unknown length
 21734 23269: contig of 1536 bp in length
 23270 23369: gap of unknown length
 23370 24647: contig of 1278 bp in length
 24648 24747: gap of unknown length
 24748 25800: contig of 1053 bp in length
 25801 25900: gap of unknown length
 25901 27017: contig of 1117 bp in length
 27018 27117: gap of unknown length
 27118 29183: contig of 2066 bp in length
 29184 29283: gap of unknown length
 29284 30315: contig of 1032 bp in length
 30316 30415: gap of unknown length
 30416 32086: contig of 1671 bp in length
 32087 32186: gap of unknown length
 32187 33395: contig of 1209 bp in length
 33396 33495: gap of unknown length
 33496 35663: contig of 2168 bp in length
 35664 35763: gap of unknown length
 35764 38044: contig of 2281 bp in length
 38045 38144: gap of unknown length
 38145 39845: contig of 1701 bp in length
 39846 39945: gap of unknown length
 39946 41983: contig of 2037 bp in length
 41983 42083: gap of unknown length
 42083 43743: contig of 1661 bp in length
 43744 43843: gap of unknown length
 43844 45904: contig of 2061 bp in length
 45905 46004: gap of unknown length
 46004 47667: contig of 1663 bp in length
 47668 47767: gap of unknown length
 47768 49833: contig of 2066 bp in length
 49834 49933: gap of unknown length
 49934 51434: contig of 1501 bp in length
 51435 51534: gap of unknown length
 51535 53778: contig of 2244 bp in length
 53779 53878: gap of unknown length
 53879 55552: contig of 1674 bp in length
 55553 55652: gap of unknown length
 55653 57670: contig of 2018 bp in length
 57671 57770: gap of unknown length
 57771 59630: contig of 1860 bp in length
 59631 59730: gap of unknown length
 59731 61147: contig of 1417 bp in length
 61148 61247: gap of unknown length
 61248 63995: contig of 2748 bp in length
 63996 64095: gap of unknown length
 64096 66294: contig of 2199 bp in length
 66295 66394: gap of unknown length
 66395 68894: contig of 2400 bp in length
 68895 71533: contig of 2639 bp in length

71534 71633: gap of unknown length
 71634 73469: contig of 1836 bp in length
 73470 73569: gap of unknown length
 73570 77054: contig of 3485 bp in length
 77055 77154: gap of unknown length
 77155 79005: contig of 1851 bp in length
 79006 79105: gap of unknown length
 79106 81294: contig of 2189 bp in length
 81295 81394: gap of unknown length
 81395 83864: contig of 2470 bp in length
 83865 83964: gap of unknown length
 83965 86039: contig of 2075 bp in length
 86040 86139: gap of unknown length
 86140 88646: contig of 2507 bp in length
 88647 88746: gap of unknown length
 88747 91823: contig of 3077 bp in length
 91824 91923: gap of unknown length
 91924 94369: contig of 2446 bp in length
 94370 94469: gap of unknown length
 94470 96886: contig of 2417 bp in length
 96887 96986: gap of unknown length
 96987 100679: contig of 3693 bp in length
 100680 100779: gap of unknown length
 100780 104171: contig of 3392 bp in length
 104172 104271: gap of unknown length
 104272 106760: contig of 2489 bp in length

Alignment Scores:
 Pred. No.: 2,3e-12 Length: 126130
 Score: 22.00 Matches: 22
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 18.80% Indels: 0
 DB: 2 Gaps: 0

US-09-897-438B-2 (1-117) x AC095877 (1-126130)

Qy 96 ValGlnPheGlnTrpLysGlnAspSerLeuArgValGlyGluValTyrGluAlaCysTrp 115
 |||||
 Db 92896 GTGCAGTTCAGTGTGAACAGGACGCTGCTGTGGTGTACGAGCCTGCTGG 92955

Qy 116 AlaLeu 117
 |||||
 Db 92956 GCCCTG 92961

RESULT 12

AC128022/c 202764 bp DNA linear HTG 19-JUL-2002
 LOCUS Rattus norvegicus clone CH230-525L20, *** SEQUENCING IN PROGRESS
 DEFINITION *** 97 unordered pieces.

ACCESSION AC128022 GI:21908605

VERSION HTG; HTGS_PHASE1.

KEYWORDS Rattus norvegicus

SOURCE Rattus norvegicus

ORGANISM Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE 1 (bases 1 to 202764)

AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,

Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Ayele,M., Banks,T.,

Barbacia,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,

Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,

Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,

Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,

Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,

Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,

Davilla,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,

Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,

Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,

Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,

Falls,T., Ferraguto,D., Flagg,N., Ford,J.J., Foster,P., Frantz,P.,

Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,

Gorrell, J. H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,
 Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L. E.,
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L. C., Lewis, L.,
 Li, J., Li, Z., Lichtarge, O., Lied, C., Liu, J., Liu, W., Loulseghe, H.,
 Lozadow, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
 Massey, E., Mawhney, E., McLeod, M. P., Meador, M., Mei, G., Metzker, M.,
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwuonu, G.,
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Pu, L. L., Quiles, M., Ren, Y.,
 Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G.,
 Scherer, S., Scott, G., Shen, H., Shooshitari, N., Sisson, I.,
 Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Usmani, K., Vasquez, L., Vera, V., Villaion, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstein, G., and Gibbs, R.
 Direct Submission
 Unpublished
 2 (bases 1 to 202764)
 Worley, K. C.
 Direct Submission
 Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: KAIY
 Center clone name: CH230-525L20
 ----- Summary Statistics
 Sequencing vector: Plasmid
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 113069 bases at least Q40
 Consensus quality: 119970 bases at least Q30
 Consensus quality: 125656 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 97 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1
 * 1244: contig of 1244 bp in length
 * 1245
 * 1344: gap of unknown length
 * 1345
 * 2346: contig of 1002 bp in length
 * 2347
 * 2446: gap of unknown length
 * 2447
 * 3813: contig of 1367 bp in length
 * 3814
 * 3913: gap of unknown length
 * 3914
 * 5101: contig of 1188 bp in length
 * 5102
 * 5201: gap of unknown length
 * 5202
 * 6724: contig of 1523 bp in length
 * 6725
 * 6824: gap of unknown length
 * 6825
 * 8048: contig of 1224 bp in length
 * 8049
 * 8148: gap of unknown length
 * 8149
 * 9212: contig of 1064 bp in length

9312: gap of unknown length
 9313
 10432: contig of 1120 bp in length
 10433
 10532: gap of unknown length
 10533
 11890: contig of 1358 bp in length
 11891
 11990: gap of unknown length
 11991
 13058: contig of 1078 bp in length
 13059
 13168: gap of unknown length
 13169
 14486: contig of 1318 bp in length
 14487
 14586: gap of unknown length
 14587
 15996: contig of 1410 bp in length
 15997
 16096: gap of unknown length
 16097
 17534: contig of 1438 bp in length
 17535
 17634: gap of unknown length
 17635
 18703: contig of 1069 bp in length
 18704
 18803: gap of unknown length
 18804
 19884: contig of 1081 bp in length
 19885
 19984: gap of unknown length
 19985
 21539: contig of 1555 bp in length
 21540
 21639: gap of unknown length
 21640
 22780: contig of 1141 bp in length
 22781
 22880: gap of unknown length
 22881
 23926: contig of 1046 bp in length
 23927
 24026: gap of unknown length
 24027
 25544: contig of 1518 bp in length
 25545
 25644: gap of unknown length
 25645
 26899: contig of 1255 bp in length
 26900
 26999: gap of unknown length
 27000
 28237: contig of 1238 bp in length
 28238
 28337: gap of unknown length
 28338
 29758: contig of 1421 bp in length
 29759
 31038: contig of 1180 bp in length
 31039
 31138: gap of unknown length
 31139
 32179: contig of 1041 bp in length
 32180
 32279: gap of unknown length
 32280
 33796: contig of 1517 bp in length
 33797
 33896: gap of unknown length
 33897
 35380: contig of 1484 bp in length
 35381
 35480: gap of unknown length
 35481
 37032: contig of 1552 bp in length
 37033
 37132: gap of unknown length
 37133
 38414: contig of 1282 bp in length
 38415
 38514: gap of unknown length
 38515
 39948: contig of 1434 bp in length
 39949
 40048: gap of unknown length
 40049
 41436: contig of 1388 bp in length
 41437
 41536: gap of unknown length
 41537
 43084: contig of 1548 bp in length
 43085
 43184: gap of unknown length
 43185
 44940: contig of 1756 bp in length
 44941
 45040: gap of unknown length
 45041
 46186: contig of 1146 bp in length
 46187
 46286: gap of unknown length
 46287
 47989: contig of 1703 bp in length
 47990
 48089: gap of unknown length
 48090
 49842: contig of 1753 bp in length
 49843
 49942: gap of unknown length
 49943
 51128: contig of 1186 bp in length
 51129
 51228: gap of unknown length
 51229
 52823: contig of 1595 bp in length
 52824
 52923: gap of unknown length
 52924
 54431: contig of 1508 bp in length
 54432
 54531: gap of unknown length
 54532
 56042: contig of 1511 bp in length
 56043
 56142: gap of unknown length
 56143
 57287: contig of 1145 bp in length
 57288
 57387: gap of unknown length
 57388
 58893: contig of 1506 bp in length
 58894
 58993: gap of unknown length
 58994
 60024: contig of 1031 bp in length
 60025
 60124: gap of unknown length
 60125
 62858: contig of 2734 bp in length
 62859
 62958: gap of unknown length

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT

* 62959	64365:	contig of 1407 bp in length
* 64366	64465:	gap of unknown length
* 64466	64434:	contig of 1969 bp in length
* 66435	66534:	gap of unknown length
* 66535	66631:	contig of 2097 bp in length
* 66632	68731:	gap of unknown length
* 68732	70939:	contig of 2108 bp in length
* 70840	70939:	gap of unknown length
* 70940	73109:	contig of 2070 bp in length
* 73110	74957:	contig of 1848 bp in length
* 74958	75057:	gap of unknown length
* 75058	76816:	contig of 1759 bp in length
* 76817	76916:	gap of unknown length
* 76917	76979:	contig of 1763 bp in length
* 78680	78779:	gap of unknown length
* 78780	80056:	contig of 1277 bp in length
* 80057	80156:	gap of unknown length
* 80157	81312:	contig of 1156 bp in length
* 81313	81412:	gap of unknown length
* 81413	82922:	contig of 1510 bp in length
* 82923	83022:	gap of unknown length
* 83023	84180:	contig of 1158 bp in length
* 84181	84280:	gap of unknown length
* 84281	85849:	contig of 1569 bp in length
* 85850	85949:	gap of unknown length
* 85950	87612:	contig of 1663 bp in length
* 87613	87712:	gap of unknown length

Alignment Scores:		
Pred. No.:	3,49e-12	Length: 202764
Score:	22.00	Matches: 22
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	18.80%	Indels: 0
DB:	2	Gaps: 0

US-09-897-438b-2 (1-117) x AC128022 (1-202764)

Qy 96 ValGlnPheGlnTnTPlysGlnAspSerLeuArgValGlyGluValTyGluAlaCysTrp 115
|||||
Db 187246 GTGACGTTCCAGTCCGAACAGGACGCTCGTGCGTGGTGGTACGAGCCCTGCTGG 187187

Qy 116 AlaLeu 117
|||||

Db 187186 GCCCTG 187181

RESULT 13
AF090441
LOCUS AF090441 10634 bp mRNA linear VRT 15-SEP-1998
DEFINITION Gallus gallus extracellular reelin (Rein) mRNA, partial cds.
ACCESSION AF090441
VERSION AF090441.1 GI:3600101

KEYWORDS Gallus gallus.

SOURCE Gallus gallus

ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.

REFERENCE 1 (bases 1 to 10634)

AUTHORS Bernier,B. and Goffinet,A.M.

TITLE Comparative study of reelin in vertebrates

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 10634)

AUTHORS Bernier,B. and Goffinet,A.M.

TITLE Direct Submission

JOURNAL Submitted (04-SEP-1998) Neurobiology, Fac.N-D de la Paix, Rue de
Bruxelles, 61, Namur B-5000, Belgium

FEATURES

source 1..10634

/organism="Gallus gallus"

/db_xref="taxon:9031"

gene <1..10634

/gene="Rein"
<1..9631
/gene="Rein"
/codon_start=2
/product="extracellular reelin"
/protein_id="AAC35559.1"
/db_xref="GI:3600102"
/translation="TGLNTTASVLQFSLGSGCRFSYSDPSITVSYSKNSADWTQ
LEKISAPSNVTIIHILYPEDAKGVNHFQKODYLHAGEVVEACWALDNLIIINAA
HRKVLEDNLDVDVTCNWLFFPGATVYKHCQSDGNSIYPHGTEGSEFATRTVDLS
TEDAQWAEFEFQKQWMDILGAVTGCTGLESSSWFLDGRKKICTPTPVDITG
YGNLRFYSGMGNGSGSHENDVILYKIEGRREHIALDTITIAAYKPSLVSVVSI
PDLQTPATKFCCLKQKSHOGRNVAWYDFHVLPLVPTVTHMIOFISNLGGCTVOPG
NSVLEFSTNHRGWSLLHTECLPEICAGPHSHFVYASENVSGNRTTTPVPAAL
TSDTRIMRQTGPIHCNMAIDNIYIGPCLKFCSCRGCTCKGCKDCPFGSGACET
ASQTFPMFISEFASRLSSYHNFYIRGAEVSGCGVLASGKALVFNKDKRQLITA
FLDSSQSRLQFTLRUGSKSVTCKAPQDQGGVLLHYSDNGITWKLHEYSYLN
HEPRIISVELPEDARIGQFRWQVYHSSQGDWAIIDEIVSYLNFNSISLDTNL
VEVTSLGYLGNVQPYCGHDMTLCFTGDSKLTSSMRYSVETSMQISAGYMFQNLV
GCGOKFTPHMDNOVKLEYSTNHGLTWHLVQEKLESPMSCOEFTSASIVHNEFTWR
RITVLLPQKTSWSSATFRWSQCYTAPDEWALDNIYIGQCPNMGSGHMGCHGVCRC
DSGFRGTECOPENPLSTVMSDEENPDVLKTEQELIGGEIVKPEBGCYVSSGSLY
FNKAGARQLYSMDLTDITWDFVOFYIQIGESSCNRREBVLQIYNSNGINWQ
LIAEMYFDSKPRFYLELPAAAKTCTFRFRWQVFGSEGVDQWAIIDITLSKQ
KHIIYPVNTLPQNFYEPAFDYPMNLQSLVWLIILANEGMTKNESFCSTPSAMLFKS
DGRFAVTRDLTLKPGYVLQFKLINIGTCTNOYSSAPVLLQYSHDAGLFWSLVKGCP
ASPTKGCESGRELSEPTVYHTGDFEDWTRITIVIPRSLAAKTRFRWIOESSHKS
VPPGIDGYIIEPCPNYCHGDCVSGVCFCDLGYTASHGTCVSNVPHSMFDFPE
RKLSPLYAKITGCGVTCGCVLSDGASLYFNGPGKREARTVPLDITNIRLVQYVQIG
SKATGSCNRPSNEGLIVQYNDNGITWHLRLREIDFMSYLEPQVVSIDLPREKTS
ATAFRWQPHGKHSQWALDDVLIIGMDSQSOTGFQDFQDVTDLQASRYRTGGQVD
IDCLSDMTALMFSENIEKPRYAEIWDHFVSASTFLFELSMGCSKPSYSHSTHLYS
LNGRDHLVTECPVPTIGCOHYTESITSERFONKRIITAYLPLTNPTRPRRW
IOYVASGVDNAIDNVLATGCPNMGSGHICDAGHCVCVCDRGGPGPYCVHNPSPV
LKDENGNLHPDLWPEVYGAERGNLNDTIKSGTALLFKEGEGLRMLVSRDLCTNTY
IQSFKIAKGTPEPERSHILQISVNGGITWHLIDFETITDVLFINVLPYTAQS
NATFRRLWQPNYSGKKEEIIIDFTIDGNLKNPILIDTDFDKPDNDWFFYPOGN
IGLYCPSYSGKGAPEEDSAMVFSNEVGEHSITTRDLSVNETIIQIEINIGCTDSS
ADPKLEFSDRLGATWHLPLCYSSSHSLSLSCSEHPSSTYTYGTQGMREVIH
FKGLHLCITRFRWYOGFYAGSOPVTAIDNVYIGPOCEEMCGHSGCINGTKICD
PGYSGTCKISTKNSDSLKDDPEGQLESDFRLLVSGGKPSKCGKSGGNSGGLFFBEG
LNMTRDLQDLQAREVQFMRLGCGKVPDRQSULSLQYSLNGLTWLSLQEFLEFS
NNSNGRYIALEIPMKARSSSTRLWQVSENGHFSYPMWIDQILIGGISSTVLED
ASCVTGSCYAIIELEYSDVIGITWHPILRDLPTNVCNRYHLRLISDTFNKWTRI
TLPLPPYTSQATFRFRWQAPDKOOWAIDNVYIGDGCIDMCSGHGCTODNKCVD
EHWGGLYCDPEPETPLTQLDNPNRSPSNQNLTVNGKLSVCGAVSGMALHFSGG
CSRMLVTDNLNTNAEFTQIFMYIGCLITPNRNQCVLLEIYVNGGITWSPLEIFPYD
QFSKPGFVNILLIPYDAKTIGTRFRWQPKHGDLDQNDWAIIDNVLSIGTDQRTVMDT
FSAPLPQHRSPPADAGPTGRIAFDMFMEDKTVNEHWFHDDCSTERCDSPDGVMY
CGSHDREYVAVTHDLPTEGIMQFKVSGCKTSEKLAQNQVHVQYSTDFGWSYL
VPQCLPADPKCSGSVSOPSVFPTKGMKVTVSLPENLVGNPVRFRYOKYSDOWAI
DNFYLGPGCLENCRGHGCDLKEQICDPCYSGPCYLTOTLTKFLKEREDNEIKPDL
WMSLBEGNCTEGCLIAEDTTLIFGGQTVQAVTQDLDRGAKFQYWGRISENNMT
TCHRPCTCRKEGVLLDSDIGGITWLLHEMDYOKYISVRHXYLILPEHALTNTLRUR
WQPFITNSIVGSPDRQWALDNLIGGAENPSQVLDTFDDEGTSHEENSSYPNA
VTAAGTCGNPSFLHYWPNKKDKTHNLSRELIQPGYMQFKIYVGCSEACSGGLH
SVMLYETKDARTDSWOLVOTHCLPSSNSIGCSPPOFHEATVYNSVNSMWRRTITQ
PDHVSSATQFRWIOKGELEKQSWAIDHVIYGEAGPKLCSGRGYCSTCAICIDGEG
QGDSCSVFSDHLPSTIKDNFESERVEITEINWETIQGGVINGCGQLAVYAHGSLYNG
CQVRQAVTKPDLSTRASKIMFVLQIGISQTSQDSCNTLDPNTVKRAVLLQSLVNNIG
TWQIAQHQPDKFIQAQRVSYNPVLEARNKGVLLRWMQPRHNGTGHQWALDHVEVVL
ISTRKONYMMNFSRQHLRHFYNRRRRSLRRYP"

BASE COUNT 2839 a 2392 c 2540 g 2862 t 1 others
ORIGIN

Alignment Scores:

Pred. No.: 5,48e-09 Length: 10634
Score: 18.00 Matches: 18
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 15.38% Indels: 0
DB: 5 Gaps: 0


```

repeat_region 40278..40311
                /rpt_family="L1"
repeat_region complement(41921..42075)
                /rpt_family="ALU"
repeat_region complement(42447..42734)
                /rpt_family="ALU"
exon 43126..43185
     /note="GRAIL prediction, score = 95"
     /evidence=not_experimental
repeat_region complement(43411..43680)
                /rpt_family="ALU"
repeat_region 45704..45991
                /rpt_family="ALU"
misc_feature 46064..46434
     /note="match to Human cDNA H60097, 5' end (NID:g1012929)"
repeat_region complement(46604..46897)
                /rpt_family="ALU"
misc_feature complement(47023..47412)
     /gene="WUGSC:H_RG249A12.1"
     /note="match to Human cDNA, 3' end H59444 (NID:g1012276)"
repeat_region 58125..58163
                /rpt_family="L1"
repeat_region 58604..58758
                /rpt_family="L1"
repeat_region 59743..59765
                /rpt_family="L1"
repeat_region 60210..60516
                /rpt_family="ALU"
exon complement(63051..63146)
     /gene="WUGSC:H_RG249A12.1"
     /note="GRAIL prediction, score = 100"
repeat_region complement(64242..64284)
                /rpt_family="L1"
repeat_region 64445..64496
                /rpt_family="L1"
repeat_region complement(67753..68011)
                /rpt_family="L1"
repeat_region 67986..68889
                /rpt_family="L1"
exon complement(70725..70841)
     /gene="WUGSC:H_RG249A12.1"
     /note="GRAIL prediction, score = 82"
repeat_region complement(71911..71947)
                /rpt_family="L1"
repeat_region complement(73631..73919)
                /rpt_family="ALU"
repeat_region 75073..75092
                /rpt_family="L1"
repeat_region 76743..82366
                /rpt_family="L1"
exon 78326..79402
     /note="GRAIL prediction, score = 83"
     /evidence=not_experimental
exon 80386..81548
     /note="GRAIL prediction, score = 82"
     /evidence=not_experimental
repeat_region complement(80962..81380)
                /rpt_family="L1"
repeat_region complement(82581..82608)
                /rpt_family="L1"
misc_feature 83877..84447

```

Alignment Scores:

```

Pred. No.:      3 69e-08      Length:      93163
Score:          18.00        Matches:      18
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:    0
Query Match:     15.38%      Indels:      0
DB:              9          Gaps:      0

```

US-09-897-438B-2 (1-117) x HSAC000121 (1-93163)

```

Oy 6 1leMethHisGlyAsnAlaValThrPheCysGluProTyGlyProArgGluLeu 23
    |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 63103 ATTATGCATGCCATGCCGTCACTTCTGTGAACCATATGCCACGAGAACTG 63050

RESULT 15
AC041023/c 185996 bp DNA linear HTG 22-MAY-2000
LOCUS Homo sapiens clone RP11-356G3, WORKING DRAFT SEQUENCE, 20 unordered
DEFINITION Pieces.
AC041023
AC041023.2 GI:8016758
VERSION HTG: HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
1 (bases 1 to 185996)
Homo sapiens, clone RP11-356G3
Unpublished
2 (bases 1 to 185996)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bedar,F.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,I., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tessfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo.A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (11-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 22, 2000 this sequence version replaced gi:7534225.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIGR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L9601
Center clone name: 356.G.3
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 175155 bases at least Q40
Consensus quality: 180623 bases at least Q30
Consensus quality: 182712 bases at least Q20
Insert size: 188000; agarose-fp
Insert size: 184096; sum-of-contigs
Quality coverage: 4.5 in Q20 bases; agarose-fp
Quality coverage: 4.6 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is

```

* arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1 2941: contig of 2941 bp in length
* 2942 3041: gap of 100 bp
* 3042 5193: contig of 2152 bp in length
* 5194 5293: gap of 100 bp
* 5294 7050: contig of 1757 bp in length
* 7051 7150: gap of 100 bp
* 7151 10452: contig of 3302 bp in length
* 10453 10552: gap of 100 bp
* 10553 13859: contig of 3307 bp in length
* 13860 13959: gap of 100 bp
* 13960 17254: contig of 3295 bp in length
* 17255 17354: gap of 100 bp
* 17355 20523: contig of 3169 bp in length
* 20524 20623: gap of 100 bp
* 20624 25148: contig of 4525 bp in length
* 25149 25248: gap of 100 bp
* 25249 29676: contig of 4428 bp in length
* 29677 29776: gap of 100 bp
* 29777 34077: contig of 4301 bp in length
* 34078 34177: gap of 100 bp
* 34178 39008: contig of 4832 bp in length
* 39010 39109: gap of 100 bp
* 39110 45564: contig of 6455 bp in length
* 45565 45664: gap of 100 bp
* 45665 54035: contig of 8371 bp in length
* 54036 54135: gap of 100 bp
* 54136 63394: contig of 9259 bp in length
* 63395 63494: gap of 100 bp
* 63495 74946: contig of 11452 bp in length
* 74947 75046: gap of 100 bp
* 75047 88353: contig of 13307 bp in length
* 88354 88453: gap of 100 bp
* 88454 104853: contig of 16400 bp in length
* 104854 104953: gap of 100 bp
* 104954 123339: contig of 18386 bp in length
* 123340 123439: gap of 100 bp
* 123440 154730: contig of 31291 bp in length
* 154731 154830: gap of 100 bp
* 154831 185996: contig of 31166 bp in length.

```

FEATURES

source

```

1. .185996
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-356G3"
/clone_lib="RPC1-11 Human Male BAC"
1. .2941
/note="assembly_fragment"
3042. .5193
/note="assembly_fragment"
5294. .7050
/note="assembly_fragment"
7151. .10452
/note="assembly_fragment"
10553. .13859
/note="assembly_fragment"
13960. .17254
/note="assembly_fragment"
17355. .20523
/note="assembly_fragment"
20624. .25148
/note="assembly_fragment"
25249. .29676
/note="assembly_fragment"
clone_end:17
vector_side:right
29777. .34077
/note="assembly_fragment"
34178. .39009

```

```

misc_feature /note="assembly_fragment"
39110. .45564
/note="assembly_fragment"
45665. .54035
/note="assembly_fragment"
54136. .63394
/note="assembly_fragment"
63495. .74946
/note="assembly_fragment"
75047. .88353
/note="assembly_fragment"
88454. .104853
/note="assembly_fragment"
104954. .123339
/note="assembly_fragment"
123440. .154730
/note="assembly_fragment"
154831. .185996
/note="assembly_fragment"
clone_end:SP6
vector_side:left
BASE COUNT 60145 a 34255 c 33646 g 56049 t 1901 others
ORIGIN

```

```

Alignment Scores:
Pred. No.: 6.78e-08 Length: 185996
Score: 18.00 Matches: 18
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 15.38% Indels: 0
DB: 2 Gaps: 0

```

US-09-897-438B-2 (1-117) x AC041023 (1-185996)

Qy 6 IleMethHisGlyAsnAlaValThrPheCysGluProTyrGlyProArgGluLeu 23

|||||
Db 166357 ATTATGCATGGCATTGCGTCACCTTCTGTGACCATATGCCCCCAGAGAACTG 166304

Search completed: November 6, 2002, 23:00:38
 Job time : 2904 secs

OM nucleic - nucleic search, using sw model

Run on: November 6, 2002, 16:42:01 ; Search time 68 Seconds
(without alignments)
1582.993 Million cell updates/sec

Title: US-09-897-438B-1
Perfect score: 351
Sequence: 1 qagcagtgtggcaccatcat.....atgaggcctgctggggcctg 351

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 441362 segs, 153338381 residues
Word size: 0

total number of hits satisfying chosen parameters: 882724

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

post-processing: Listing first 45 summaries.

```
Database : Issued_Patents_NA:*
1: /cgn2.6/ptodata/1/lna/5A_COMB.seq:*
2: /cgn2.6/ptodata/1/lna/5B_COMB.seq:*
3: /cgn2.6/ptodata/1/lna/6A_COMB.seq:*
4: /cgn2.6/ptodata/1/lna/6B_COMB.seq:*
5: /cgn2.6/ptodata/1/lna/PCTUS_COMB.seq
6: /cgn2.6/ptodata/1/lna/backfiles.seq
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description
	1	351	100.0	11673	4	US-09	334-220-3	Sequence 3, Appli
	2	26	7.4	11580	4	US-09	334-220-4	Sequence 4, Appli
	3	18	5.1	1074	2	US-08	627-151A-15	Sequence 15, Appl
	4	18	5.1	1404	6	5171840	-8	Patent No. 5171840
	5	18	5.1	1404	6	5480796	-8	Patent No. 5480796
	6	18	5.1	1486	4	US-08	795-473B-3	Sequence 3, Appli
	7	18	5.1	1486	4	US-09	439-856-3	Sequence 3, Appli
	8	18	5.1	2061	6	5171840	-1	Patent No. 5171840
	9	18	5.1	2061	6	5480796	-1	Patent No. 5480796
	10	18	5.1	3319	4	US-08	795-473B-2	Sequence 2, Appli
	11	18	5.1	3319	4	US-09	439-856-2	Sequence 2, Appli
	12	17	4.8	47	2	US-08	642-684-13	Sequence 13, Appl
C	13	17	4.8	1434	4	US-09	480-921B-9	Sequence 9, Appli
C	14	17	4.8	1434	4	US-09	480-921B-28	Sequence 28, Appl
	15	17	4.8	1488	4	US-09	108-010B-11	Sequence 11, Appl
C	16	17	4.8	1732	4	US-09	449-335-1	Sequence 1, Appli
C	17	17	4.8	1732	4	US-09	449-335-5	Sequence 5, Appli
	18	17	4.8	1770	2	US-08	642-684-3	Sequence 3, Appli
	19	17	4.8	2935	4	US-09	480-921B-27	Sequence 27, Appl
	20	17	4.8	6386	2	US-08	483-376-1	Sequence 1, Appli
	21	17	4.8	8050	4	US-09	491-362-11	Sequence 11, Appl
	22	17	4.8	8050	4	US-09	874-562-11	Sequence 11, Appl
C	23	16	4.6	2892	4	US-08	981-392-11	Sequence 11, Appl
C	24	16	4.6	2857	4	US-08	981-392-4	Sequence 4, Appli
C	25	16	4.6	3600	4	US-09	657-042A-3	Sequence 3, Appli
C	26	16	4.6	4181	4	US-09	393-569-1	Sequence 1, Appli
C	27	16	4.6	5160	3	US-08	895-601-1	Sequence 1, Appli

ALIGNMENTS

28	16	4.6	5430	3	US-09-012-515A-11	Sequence 11, Appl
29	16	4.6	5430	3	US-08-360-144A-11	Sequence 11, Appl
30	16	4.6	5430	4	US-09-012-504A-11	Sequence 11, Appl
31	16	4.6	6911	1	US-08-311-174-4	Sequence 1, Appl
32	16	4.6	7653	4	US-08-471-112A-1	Sequence 1, Appl
33	16	4.6	7824	5	PCT-US95-06722-11	Sequence 11, Appl
34	16	4.6	8948	4	US-09-643-597-119	Sequence 119, Appl
35	15	4.3	81	4	US-09-133-321-7	Sequence 7, Appl
36	15	4.3	307	4	US-09-615-192A-165	Sequence 165, Appl
37	15	4.3	358	2	US-08-454-557C-15	Sequence 15, Appl
38	15	4.3	358	2	US-08-340-426D-15	Sequence 15, Appl
39	15	4.3	358	2	US-08-450-673C-15	Sequence 15, Appl
40	15	4.3	358	5	PCT-US95-17111A-15	Sequence 15, Appl
41	15	4.3	463	1	US-07-753-101A-67	Sequence 67, Appl
42	15	4.3	531	2	US-08-809-267-9	Sequence 9, Appl
43	15	4.3	531	5	PCT-US95-13662A-9	Sequence 9, Appl
44	15	4.3	557	4	US-09-404-879A-88	Sequence 88, Appl
45	15	4.3	557	4	US-09-404-879A-87	Sequence 87, Appl

RESULT 1
US-09-334-220-3
Sequence 3. Application US/09334220

patent NO. 6323177
; GENERAL INFORMATION:
APPLICANT: St. Jude's Children's Research Hospital

APPLICANT: CURRAN, Thomas
APPLICANT: D'Arcangelo, Gabriella
TITLE OF INVENTION: INTERACTION OF REELIN WITH VERY LOW DENSITY LIPOPROTEIN (VLDL) RECEPTOR FOR SCREENING AND THERAPIES

1. TYPE OF INFORMATION - 2427/0E704

: CURRENT APPLICATION NUMBER: US/09/334,220

CURRENT FILING DATE: 1999-06-16

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: Fast

SEQ ID NO. 3

```

; LENGTH: 11
; TYPE: DNA

```

TYPE: DNA
ORGANISM: *Mus musculus*

US-09-334-220-3

Query Match	100.0%	Score 351;	DB 4;	Length 11673;
Best Local Similarity	100.0%	Pred. NO. 1.4e-174;		
Matches 351;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

[illegible]

QY 301 AACAGGACAGCCTGCGAGTGGGTGAGGTGTATCAGGCCCTGCTGGGCCCTG 351

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 6, 2002, 15:27:06 ; Search time 300 Seconds
(without alignments)
2634.838 Million cell updates/sec

Title: US-09-897-438B-1
Perfect score: 351
Sequence: 1 gaggcgtgtggcaccatcat.....atgaggcctgtggccttg 351

Scoring table:
OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

W size : 0
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	351	100.0	351	ABL40165	Mouse reelin prote
2	351	100.0	2745	AAD22754	Mus musculus trunc
3	351	100.0	11673	ABA92603	Mouse reelin encod
4	351	100.0	11673	AB199284	Mouse ischaemic co
5	26	7.4	11580	ABN96939	Gene #3437 used to
6	26	7.4	11580	ABA92604	Human reelin encod
7	26	7.4	11632	AAS89484	DNA encoding novel
8	22	6.3	22	ABL40169	Mouse reelin prote
9	21	6.0	21	ABL40173	Mouse reelin prote

10	20	5.7	20	24	AAD22777	Mouse truncated re
11	19	5.4	4883	22	AAH19492	Human coding seque
12	19	5.4	7215	22	AAH19497	Human coding seque
13	19	5.4	7215	22	AAH19498	Human coding seque
14	18	5.1	18	24	ABL40174	Mouse reelin prote
15	18	5.1	975	21	AAH70701	Human interleukin
16	18	5.1	1035	21	AAZ40288	SR345 coding seque
17	18	5.1	1074	18	AAV04440	Interleukin 6 rece
18	18	5.1	1260	20	AAZ09202	Human IL-6 recepto
19	18	5.1	1486	19	AAV60296	Human interleukin-
20	18	5.1	1545	21	AAH70763	IL-6R/IL-6 fusion
21	18	5.1	1627	18	AAH97848	Human fusion polyp
22	22	5.1	2061	10	AAH90340	Sequence encoding
23	18	5.1	2066	14	AAQ41746	IL-6 receptor codi
24	18	5.1	2087	10	AAH90847	DNA contg. region
25	18	5.1	3004	21	AAZ33369	Human secreted pro
26	18	5.1	3319	17	AAH31441	Interleukin-6 rece
27	18	5.1	3319	19	AAV60295	Human interleukin-
28	18	5.1	3319	21	AAH21364	Human low adenosin
29	18	5.1	3319	21	AAH35242	Human adenosine re
30	18	5.1	3319	24	ABK84527	Human cDNA differe
31	18	5.1	3477	21	AAA09047	Fusion polypeptide
32	18	5.1	3507	21	AAA09046	Fusion polypeptide
33	18	5.1	4513	21	AAH21365	Human low adenosin
34	18	5.1	4873	21	AAH35243	Human adenosine re
c 35	18	5.1	9720	24	ABL56176	Hordeum vulgare va
c 36	18	5.1	198285	24	ABK84699	Human cDNA differe
c 37	18	5.1	198285	24	ABN97319	Gene #3817 used to
38	17	4.8	22	21	AAH63485	Beta-actin gene PC
39	17	4.8	47	18	AAH90826	Bacillus stearothe
c 40	17	4.8	325	22	AAH36445	Human musculoskele
c 41	17	4.8	395	18	AAH13260	N. tabacum strain
c 42	17	4.8	429	24	ABN95973	Gene #2471 used to
43	17	4.8	559	22	ABA63804	Human foetal liver
44	17	4.8	559	22	ABA30992	Probe #9458 for ge
45	17	4.8	559	22	AAH12314	Human brain expres

ALIGNMENTS

RESULT 1
ABL40165
ID ABL40165 standard; DNA; 351 BP.
XX
AC ABL40165;
XX
DT 21-MAY-2002 (first entry)
XX
DE Mouse reelin protein CR-50 epitope region encoding DNA SEQ ID NO:1.
DE DE
KW Mouse; reelin protein CR-50 epitope region; elucidation; neuron;
KW cerebral disturbance; reelin protein; neuroprotective; gene; ds.
XX
OS Mus musculus.
XX
FH key Location/Qualifiers
FT CDS 1..351
FT /tag= a
FT /product= reelin protein CR-50 epitope region
FT /note= "no start or stop codons given"

JP2002017361-A.

22-JAN-2002.

04-JUL-2000; 2000JP-0202801.

04-JUL-2000; 2000JP-0202801.

(RIKA) RIKAGAKU KENKYUSHO.

DR WPI; 2002-221707/28.
XX P-PSDB; ABB06244.
PT Reelin protein CR-50 epitope region, useful for diagnosis and treatment
PT of cerebral disturbance
XX
PS
XX Claim 5; Page 11; 16pp; Japanese.
CC The present sequence encodes the mouse reelin protein CR-50 epitope
CC region, which contains the CR-50 antibody recognition site and is free
CC from F-spondin domains and repetitive sites. Also described are: (1) an
CC expression vector comprising a polynucleotide encoding a reelin protein
CC epitope region; (2) host cells with transfected the expression vector;
CC (3) polypeptides prepared by culture of the host cells; and (4)
CC polynucleotides comprising the 351 base sequence given in ABL40165 which
CC encodes the 117 amino acid sequence given in ABB06244; and (5) use of
CC the polynucleotide for diagnosis and/or treatment of diseases caused by
CC abnormal positioning of neural cells, and stimulation of association of
CC reelin protein. The mouse reelin protein CR-50 epitope region has
CC neuroprotective activity, and can be used in the diagnosis and treatment
CC of cerebral disturbance due to an abnormal reelin gene and positioning
CC of neurons.
XX
SQ Sequence 351 BP; 86 A; 98 C; 92 G; 75 T; 0 other;
Query Match 100.0%; Score 351; DB 24; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.6e-166;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGCAGTGTGGCACCACATCATGCAATGCTGTCACTTCTGTGAGCCGTACGGCCCT 60
Db 1 GAGCAGTGTGGCACCACATCATGCAATGCTGTCACTTCTGTGAGCCGTACGGCCCT 60
QY 61 CGAGAGTGTGGCACCACATCATGCAATGCTGTCACTTCTGTGAGCCGTACGGCCCT 120
Db 61 CGAGAGTGTGGCACCACATCATGCAATGCTGTCACTTCTGTGAGCCGTACGGCCCT 120
QY 121 GGGTCAGGATCATGTCGATTTAGTTACTTCTGACCCAGCATCTGTCTATACGCCAAG 180
Db 121 GGGTCAGGATCATGTCGATTTAGTTACTTCTGACCCAGCATCTGTCTATACGCCAAG 180
QY 181 AACATACCCGCTGATGGATTGAGTTCAGTGGAGAAATTTAGAGCCCTTCCAAATGTGAGCACA 240
Db 181 AACATACCCGCTGATGGATTGAGTTCAGTGGAGAAATTTAGAGCCCTTCCAAATGTGAGCACA 240
QY 241 GTCATCCACATCTGTACTCCCGAGGAGAACCAAGGGGAGAGCGGTTCAGTTCCAGTGG 300
Db 241 GTCATCCACATCTGTACTCCCGAGGAGAACCAAGGGGAGAGCGGTTCAGTTCCAGTGG 300
QY 301 AACAGGAGACGCTGCGAGTGGGTGAGTGTATGAGGCGCTGCTGGGCCCTG 351
Db 301 AACAGGAGACGCTGCGAGTGGGTGAGTGTATGAGGCGCTGCTGGGCCCTG 351
RESULT 2
AAD22754
ID AAD22754 standard; cDNA; 2745 BP.
XX
AC AAD22754;
XX
DT 26-FEB-2002 (first entry)
XX
XX Mus musculus truncated reelin cDNA.
DE
XX Mouse; reelin; F-spondin domain; CR-50 epitope; gene therapy; agyria;
KW polymicrogyria; ectopic gray matter; ss.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 283..2052
FT /*tag= a
FT /product= "Mouse truncated reelin protein"

FT sig_peptide 283..363
FT /*tag= b
FT mat_peptide 364..2049
FT /*tag= c
FT /product= "Mature truncated reelin protein"
FT 284..849
FT /*tag= d
FT /note= "Encodes F-spondin domain"
FT 970..1320
FT /*tag= e
FT /note= "Encodes CR-50 epitope region"
XX
XX EP1149844-A2.
XX 31-OCT-2001.
XX 11-APR-2001; 2001BP-0303411.
XX 11-APR-2000; 2000JP-0109954.
XX
XX (PKE) RIKEN KK.
XX Mikoshiba K, Tabata H, Nakajima K;
XX WPI; 2002-019320/03.
XX P-PSDB; AAE13606.
XX Novel truncated Reelin protein containing F-spondin domain and CR-50
XX recognition site of Reelin protein, but not having Reelin repeat site,
XX useful to treat diseases including agyria due to abnormal neuron
XX alignment
XX
PS Claim 10; Page 20-26; 47pp; English.
CC The invention relates to a truncated Reelin protein comprising a
CC F-spondin domain and a CR-50 recognition site but no reelin protein
CC repeat site. Reelin is an essential molecule in developing a normal
CC laminated structure of cerebrum. The truncated reelin protein and its
CC DNA are useful for treating diseases including agyria, polymicrogyria,
CC and ectopic gray matter due to abnormal neuronal alignment. Truncated
CC reelin protein DNA is useful in gene therapy. The present sequence is
CC a cDNA encoding Mus musculus truncated reelin protein.
XX
SQ Sequence 2745 BP; 661 A; 716 C; 714 G; 654 T; 0 other;
Query Match 100.0%; Score 351; DB 24; Length 2745;
Best Local Similarity 100.0%; Pred. No. 1.5e-166;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGCAGTGTGGCACCACATCATGCAATGCTGTCACTTCTGTGAGCCGTACGGCCCT 60
Db 970 GAGCAGTGTGGCACCACATCATGCAATGCTGTCACTTCTGTGAGCCGTACGGCCCT 1029
QY 61 CGAGAGTGTGGCACCACATCATGCAATGCTGTCACTTCTGTGAGCCGTACGGCCCT 120
Db 1030 CGAGAGTGTGGCACCACATCATGCAATGCTGTCACTTCTGTGAGCCGTACGGCCCT 1089
QY 121 GGGTCAGGATCATGTCGATTTAGTTACTTCTGACCCAGCATCTGTCTATACGCCAAG 180
Db 1090 GGGTCAGGATCATGTCGATTTAGTTACTTCTGACCCAGCATCTGTCTATACGCCAAG 1149
QY 181 AACATACCCGCTGATGGATTGAGTTCAGTGGAGAAATTTAGAGCCCTTCCAAATGTGAGCACA 240
Db 1150 AACATACCCGCTGATGGATTGAGTTCAGTGGAGAAATTTAGAGCCCTTCCAAATGTGAGCACA 1209
QY 241 GTCATCCACATCTGTACTCCCGAGGAGAACCAAGGGGAGAGCGGTTCAGTTCCAGTGG 300
Db 1210 GTCATCCACATCTGTACTCCCGAGGAGAACCAAGGGGAGAGCGGTTCAGTTCCAGTGG 1269
QY 301 AACAGGAGACGCTGCGAGTGGGTGAGTGTATGAGGCGCTGCTGGGCCCTG 351
Db 1270 AACAGGAGACGCTGCGAGTGGGTGAGTGTATGAGGCGCTGCTGGGCCCTG 1320

US Designate

RESULT: 2
24926f

THIS PAGE BLANK (USPTO)

QY	1	GAGCAGTGTGGCACCACCATCATGCGATGGCAATCTGTCACTTCTGTGTGAGCCGTACGGCCCT	60
Db	970	GAGCAGTGTGGCACCACCATCATGCGATGGCAATCTGTCACTTCTGTGTGAGCCGTACGGCCCT	1029
QY	61	CGAGAGCTGACCACACACATGCGCTGAACACACACACAGACATCTGCTCCAGTTCCTCATT	120
Db	1030	CGAGAGCTGACCACACACATGCGCTGAACACACACACAGACATCTGCTCCAGTTCCTCATT	1089
QY	121	GGTCCAGGATCATGTCGATTTTACTTCTGACCCCGAGCATCACTGTGTCTCATACGCCAAG	180
Db	1090	GGTCCAGGATCATGTCGATTTTACTTCTGACCCCGAGCATCACTGTGTCTCATACGCCAAG	1149
QY	181	AACAATACCGCTGATTGGATTTCAGCTGGAGAAAATTAGAGCCCTTCCATGTGAGGCACA	240
Db	1150	AACAATACCGCTGATTGGATTTCAGCTGGAGAAAATTAGAGCCCTTCCATGTGAGGCACA	1209
QY	241	GTATCCACATCCTGTACCTCCCGGAGGAAGCAAAAGGGAGAGCGTGCAGTTCCTCAGTGG	300
Db	1210	GTATCCACATCCTGTACCTCCCGGAGGAAGCAAAAGGGAGAGCGTGCAGTTCCTCAGTGG	1269
QY	301	AAACAGACAGCCTCGGAGTGGGTGAGTGTATGAGGCCCTCTGGGCCCCG	351
Db	1270	AAACAGACAGCCTCGGAGTGGGTGAGTGTATGAGGCCCTCTGGGCCCCG	1320
RESULT 4			
ABI99284			
ID	ABI99284	standard; cDNA; 11673 BP.	
AC	XX	ABI99284;	
CC	XX		
GG	XX		
TT	XX		
DD	XX	07-MAR-2002 (first entry)	
EE	XX	Mouse ischaemic condition related cDNA sequence	SEQ ID NO:128.
FF	XX		
KK	Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;		
WW	vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.		

18-MAY-2001; 2001WO-JP04192.
18-MAY-2000; 2000JP-0145977.
(YUNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
WPI; 2002-034733/04.
P-PSDB; ABB57065.
Examining the ischemic condition (e.g. occlusive ischemia) by measuring
expression levels of particular genes defined in the specification or
by determining the expression profile of a gene group comprising these
genes -
Claim 2; Page 362-385; 2690pp; English.
The present invention describes a method for examining ischaemic
conditions, comprising measuring the expression levels of particular
genes (I) in a test sample or determining the expression profile of a
gene group in the sample comprising genes selected from (I). The method
is useful for examining the ischaemic condition (e.g. compressive
ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
expression levels of particular genes (ABI99202 to ABI99912, encoding
the protein sequences in ABB57020 to ABB57374) or by determining the
expression profile of a gene group comprising these genes. The
expression levels or expression profiles produced by these genes are
used as an indicator when screening for ischaemic condition-improving

is useful for examining the ischaemic condition (e.g. compressive ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring expression levels of particular genes (AB199202 to AB199912, encoding the protein sequences in AB557020 to AB557374) or by determining the expression profile of a gene group comprising these genes. The expression levels or expression profiles produced by these genes are used as an indicator when screening for ischaemic condition-improving

THIS PAGE BLANK (USPTO)

progression of liver cancer, hepatocellular carcinoma or metastatic liver cancer in a patient, and differentiating metastatic liver cancer from hepatocellular carcinoma in a patient, involving detecting the level of expression of two or more genes represented in ABN3503-ABN97455 in a tissue sample. The method of the invention has hepatotropic, and cytototoxic activity. The method is useful for diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma and metastatic liver carcinoma in a patient. The method is useful for identifying expression profiles which serve as useful diagnostic markers as well as markers that can be used to monitor disease states, disease progression, drug toxicity, drug efficacy and drug metabolism.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

cytotoxic activity. The method is useful for diagnosing a
the progression of liver cancer, hepatocellular carcinoma a
liver carcinoma in a patient. The method is useful for iden
expression profiles which serve as useful diagnostic markers
markers that can be used to monitor disease states, disease
drug toxicity, drug efficacy and drug metabolism.
Note: The sequence data for this patent did not form part o
specification, but was obtained in electronic format direct
at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).
Sequence 11580 BP: 3014 A; 2696 C; 2753 G; 3116 T; 1 other:
XX

markers that can be used to monitor drug toxicity, drug efficacy and drug

CC-0. Note: The sequence data for this patent did not specify the exact sequence, but was obtained in electronic form from the patent office.

CC-0 at: ftp.wipo.int/pub/published_po
XX

Sequence 11580 BP; 3014 A; 2696 C; 2753 G; 3116 T; 1 other;

Query Match 7.4%; Score 26; DB 24; Length 11580;
Best Local Similarity 100.0%; Pred No 0.0041.

```
Matches 26; Conservative 0; Mismatches 0;
```

QY 208 · GAGAAATTAGAGCCCTTCCAATGT 233

Db 1067 GAGAAAATTAGAGCCCCCTTCCAATGT 1092

RESULT 6

ABA92604
ID ABA92604 standard; cdNA: 11580 BP.

XX
AC ABA92604;

21-MAR-2002 (first entry)

XX	DE	Human reelin encoding cDNA	SEQ ID NO
1	1	1	1
2	2	2	2
3	3	3	3
4	4	4	4
5	5	5	5
6	6	6	6
7	7	7	7
8	8	8	8
9	9	9	9
10	10	10	10
11	11	11	11
12	12	12	12
13	13	13	13
14	14	14	14
15	15	15	15
16	16	16	16
17	17	17	17
18	18	18	18
19	19	19	19
20	20	20	20
21	21	21	21
22	22	22	22
23	23	23	23
24	24	24	24
25	25	25	25
26	26	26	26
27	27	27	27
28	28	28	28
29	29	29	29
30	30	30	30
31	31	31	31
32	32	32	32
33	33	33	33
34	34	34	34
35	35	35	35
36	36	36	36
37	37	37	37
38	38	38	38
39	39	39	39
40	40	40	40
41	41	41	41
42	42	42	42
43	43	43	43
44	44	44	44
45	45	45	45
46	46	46	46
47	47	47	47
48	48	48	48
49	49	49	49
50	50	50	50
51	51	51	51
52	52	52	52
53	53	53	53
54	54	54	54
55	55	55	55
56	56	56	56
57	57	57	57
58	58	58	58
59	59	59	59
60	60	60	60
61	61	61	61
62	62	62	62
63	63	63	63
64	64	64	64
65	65	65	65
66	66	66	66
67	67	67	67
68	68	68	68
69	69	69	69
70	70	70	70
71	71	71	71
72	72	72	72
73	73	73	73
74	74	74	74
75	75	75	75
76	76	76	76
77	77	77	77
78	78	78	78
79	79	79	79
80	80	80	80
81	81	81	81
82	82	82	82
83	83	83	83
84	84	84	84
85	85	85	85
86	86	86	86
87	87	87	87
88	88	88	88
89	89	89	89
90	90	90	90
91	91	91	91
92	92	92	92
93	93	93	93
94	94	94	94
95	95	95	95
96	96	96	96
97	97	97	97
98	98	98	98
99	99	99	99
100	100	100	100

Human; reelin; low density lipoprotein receptor; LDLR; neur

KW extracellular glycoprotein; nontropic; antilipase

KW neurodegenerative disorder; neuronal regeneration

Lipid metabolism disease; memory; d
KW
XX

US
XX
Homo sapiens.

EFH	key	Location/Qualifiers
FT	CDS	176..10558

```

/*tag= a
/product="human reelin"

```

XX
PN US6323177-B1.

XX
PD 27-NOV-2001.

16-JUN-1999; 99US-0334220.

16-JUN-1999; 99US-0334220.

PA (SJUD-) ST JUDE CHILDREN'S RES HOSP

XX
PI Curran T, D'Arcangelo G;

XX
DR
WPI; 2002-096596/13.

OR P-PSDB; ABB05007.
XX

Novel composition useful for screening binding to low density lipoprotein

PT Reelin polypeptide and low density lipoprotein receptor -
XX

Example 1; Column 75-84; 45pp; Engl: XX

CCC The present invention describes a composition (I) comprising isolated reelin protein (II) bound to an isolated low density lipoprotein receptor (III) or a fragment thereof.

receptor (LDLR) (III). (II) is an ex-

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.
OM nucleic - nucleic search, using sw model
Run on: November 6, 2002, 16:41:41 ; Search time: 2213 Seconds
(without alignments)
2568.737 Million cell updates/sec

Title: US-09-897-438B-1
Perfect score: 351
Sequence: 1 gagcagtggtgaccatcat.....atgaggcctgctgggcccctg 351

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estinu:*
- 4: em_estnu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_hic:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_hic:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estfun:*
- 16: em_estom:*
- 17: gb_gss:*
- 18: em_gss_hum:*
- 19: em_gss_inv:*
- 20: em_gss_pln:*
- 21: em_gss_vrt:*
- 22: em_gss_fun:*
- 23: em_gss_mam:*
- 24: em_gss_mus:*
- 25: em_gss_other:*
- 26: em_gss_pro:*
- 27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Match	Length	Description
C 1	134	38.2	600	12 BG803882
C 2	94	26.8	635	10 BB248113
C 3	26	7.4	315	12 BE938667
C 4	22	6.3	568	12 BF387628
C 5	20	5.7	431	10 BB819780
C 6	19	5.4	404	14 H58973

C 7	19	5.4	416	9 AI400688
C 8	19	5.4	525	10 AW178252
C 9	19	5.4	539	10 AW503344
C 10	19	5.4	741	10 BE288244
C 11	19	5.4	784	13 BM051508
C 12	19	5.4	845	12 BE974073
C 13	19	5.4	899	12 BF680728
C 14	19	5.4	1007	12 BG484905
C 15	19	5.4	1575	13 BM480049
C 16	19	5.1	102	17 BH870562
C 17	18	5.1	125	17 BH779420
C 18	18	5.1	164	17 BH124575
C 19	18	5.1	187	10 BB118397
C 20	18	5.1	304	10 BB605282
C 21	18	5.1	306	17 AZ765799
C 22	18	5.1	320	12 BF748992
C 23	18	5.1	377	17 AZ816986
C 24	18	5.1	379	9 AA068201
C 25	18	5.1	379	9 AA583727
C 26	18	5.1	384	9 AA794374
C 27	18	5.1	395	17 AZ106948
C 28	18	5.1	399	14 BQ329315
C 29	18	5.1	414	9 AA502373
C 30	18	5.1	449	12 BF706216
C 31	18	5.1	452	17 BH775447
C 32	18	5.1	458	14 BQ335482
C 33	18	5.1	475	12 BF390018
C 34	18	5.1	477	10 AV918544
C 35	18	5.1	480	17 BH782630
C 36	18	5.1	488	17 BH784939
C 37	18	5.1	490	17 BH882572
C 38	18	5.1	493	17 AQ215835
C 39	18	5.1	502	12 BH873530
C 40	18	5.1	512	12 BG737680
C 41	18	5.1	516	10 AV947184
C 42	18	5.1	548	12 BF631988
C 43	18	5.1	574	13 BJ482511
C 44	18	5.1	582	17 BH873271
C 45	18	5.1	627	12 BG446777

ALIGNMENTS

RESULT 1
BG803882/c
LOCUS 0243-03 Mouse El4.5 retina lambda ZAP II Library Mus musculus cDNA, EST 20-DEC-2001
DEFINITION mRNA sequence.
ACCESSION BG803882.1 GI:17950794
VERSION BG803882.1
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 600)
AUTHORS Mu, X., Zhao, S., Pershad, R., Hsieh, T.-F., Scarpa, A., Wang, S.W., White, R.A., Beremand, P.D., Thomas, T.L., Gan, L., and Klein, W.H.
TITLE Gene expression in the developing mouse retina by EST sequencing and microarray analysis
JOURNAL Nucleic Acids Res. 29 (24), 4983-4993 (2001)
MEDLINE 21671825
COMMENT Contact: Klein WH
Department of Biochemistry and Molecular Biology
University of Texas M.D. Anderson Cancer Center
Box 117, 1515 Holcombe Blvd., Houston, TX 77030, USA
Tel: 713 792 3646
Fax: 713 790 0329.
LOCATION/Qualifiers
1..600
/organism="Mus musculus"
/db_xref="taxon:10090"

AI400688 tg93e01.x
AW178252 MRO-HT006
AW503344 UI-HF-BNO
BE288244 601095415
BM051508 603638192
BE974073 601680409
BF680728 602155347
BG484905 602503910
BM480049 AGENCOURT
BH870562 hm64b04.b
BH779420 fzmb014f0
BH124575 RPCI-24-3
BB118397 BB118397
BB605282 BB605282
AZ765799 IM0562M21
BF748992 MR2-BN038
AZ816986 2M0085O21
AA068201 mm48b04.r
AA583727 nm63e04.s
AA794374 vU68a07.f
AZ106948 RPCI-23-4
BQ329315 RC6-EN008
AA502373 ne27e09.s
BF706216 280592.MA
BH775447 fzmb011f0
BQ335482 FM4-MT043
BF390018 UI-R-BS2-
AV918544 AV918544
BH782630 fzmb011f0
BH784939 fzmb013f0
BH882572 hw38c01.g
AQ215835 HS_3252.A
BH873530 hp45e10.g
BG737680 f094e09.y
AV947184 AV947184
BF631988 PM3-HT090
BJ482511 BJ482511
BH873271 hp43f10.b
BG446777 GA_EB003

```

/clone_lib="Mouse E14.5 retina lambda ZAP II Library"
/tissue_type="neural retina"
/dev_stage="embryonic day 14.5 post-fertilization"
/notes="Vector: pAMP10 (Gibco); Cloned unidirectionally.
Primer: Oligo dT. RNA Isolation: cytoplasmic RNA preps
(Mannatis); Cloning Technique: CUA Cloning (CloneAmp,
Life Technologies); Average insert size: 1.8 kb;
Insertion site: TACGTCCACTGAATTCGTAGTG--->. Other
information regarding entire library may be found at
http://ppa.swmed.edu/bata/Libraries/microarray_cdna_librar
ies.htm."

```

```

BASE COUNT      127 a  167 c  152 g  152 t      2 others
ORIGIN
Query Match      38.2%; Score 134; DB 12; Length 600;
Best Local Similarity 100.0%; Pred. No. 2.5e-62;
Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 218 GAGCCCTTCCAATGTGAGCACAGTCATCCACATCTCTGTACCTCCCGAGGAAGCAAG 277
Db 476 GAGCCCTTCCAATGTGAGCACAGTCATCCACATCTCTGTACCTCCCGAGGAAGCAAG 417
QY 278 GGGAGAGCGTGCAGTTCACGTGGAACAGACAGCCCTGCGAGTGGGTGAGGTATCAGG 337
Db 416 GGGAGAGCGTGCAGTTCACGTGGAACAGACAGCCCTGCGAGTGGGTGAGGTATCAGG 357
QY 338 CCTGCTGGGCCCTG 351
Db 356 CCTGCTGGGCCCTG 343

```

```

RESULT 2
BB248113
LOCUS
DEFINITION
BB248113 635 bp mRNA linear EST 23-OCT-2001
musculus cDNA clone A730023J04 3', mRNA sequence.

```

```

ACCESSION
BB248113
VERSION
KEYWORDS
SOURCE
EST.
GI:16355611

```

```

ORGANISM
Mus musculus
house mouse.

```

```

REFERENCE
AUTHORS
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Sasaki
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)

```

```

TITLE
JOURNAL
COMMENT
On Jul 6, 2000 this sequence version replaced gi:8940859.

```

```

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/

```

```

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura
S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format

```

```

sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

```

FEATURES

source

```

Location/Qualifiers
1..635
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="A730023J04"
/clone_lib="RIKEN full-length enriched, 7 days neonate
cerebellum"
/tissue_type="cerebellum"
/dev_stage="7 days neonate"
/lab_host="DH10B"
/notes="Site_1: Salt; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGGATCCAGAGCTCTTTTCTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 20.0 and subtraction to Rot = 459.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGAGAGAGATTCGAGTTAATAATATCCCTCCCTCCCTCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I."

```

```

BASE COUNT      182 a  122 c  118 g  213 t
ORIGIN
Query Match      26.8%; Score 94; DB 10; Length 635;
Best Local Similarity 100.0%; Pred. No. 2.3e-40;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 29 ATGCTGTACCTTCTGTGAGCGCTACGGCCTCGAGAGTCGACACACATGCTGAACA 88
Db 1 ATGCTGTACCTTCTGTGAGCGCTACGGCCTCGAGAGTCGACACACATGCTGAACA 60

```

```

QY 89 CAACAACAGCATCTGTCCTCCAGTTTTCATTGG 122
Db 61 CAACAACAGCATCTGTCCTCCAGTTTTCATTGG 94

```

RESULT 3

```

BE938667/c

```

```

LOCUS
BE938667

```

```

DEFINITION
QV0-TN0084--180800-342-a08 TN0084 Homo sapiens cDNA, mRNA sequence.

```

```

ACCESSION
BE938667

```

```

VERSION
BE938667.1 GI:10466224

```

```

KEYWORDS
EST.

```

```

SOURCE
human.

```

```

ORGANISM
Homo sapiens

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```


REFERENCE
AUTHORS

1 (bases 1 to 315)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL
MEDLINE
COMMENT

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

{http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-QV0-7N0084-180}

800-342-a086t3-2000-08-18&t4=1

Seq primer: puc 18 forward

High quality sequence stop: 315.

FEATURES

source

Location/Qualifiers

1..315

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="TN0084"

/dev_stage="Adult"

/note="Organ: testis_normal; Vector: puc18; Site_1: Sma1;

Site_2: Sma1; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

84 a 54 c 78 g 89 t

BASE COUNT

ORIGIN

Query Match 7.4%; Score 26; DB 12; Length 315;

Best Local Similarity 100.0%; Pred. No. 0.004;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 208

GAGAAATTAGAGCCCTTCCAAATGT 233

|||||

Db 178 GAGAAATTAGAGCCCTTCCAAATGT 153

RESULT 4

BF387628/c

LOCUS

BF387628 568 bp mRNA linear EST 27-NOV-2000

UI-R-CA1-bbs-e-06-0-UI.s1 UI-R-CA1 Rattus norvegicus cDNA clone

UI-R-CA1-bbs-e-06-0-UI 3', mRNA sequence.

ACCESSION

BF387628

VERSION

BF387628.1

KEYWORDS

EST.

SOURCE

ORGANISM

Rattus norvegicus

Norway rat.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 568)

REFERENCE

AUTHORS

Bonaldo, M.F., Lennon, G., and Soares, M.B.

TITLE

Normalization and subtraction: Two approaches to facilitate gene

discovery

JOURNAL

Genome Res. 6 (9), 791-806 (1996)

MEDLINE

97044477

COMMENT

Contact: Soares, MB

Program for Rat Gene Discovery and Mapping

University of Iowa

451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu

The sequence contained an oligo-dT track that was present in the

oligonucleotide that was used to prime the synthesis of first

strand cDNA and therefore this may represent a bonafide poly A

tail. The sequence tag present in the cDNA between the NotI site

and the oligo-dT track served to identify it as a clone from the

normalized pons library cDNA Library Preparation: M.B. Soares Lab

Clone distribution: clones will be available through Research

Genetics (www.resgen.com)

Seq primer: M13 Forward

POLYA=Yes

FEATURES

source

Location/Qualifiers

1..568

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-CA1-bbs-e-06-0-UI"

/clone_lib="UI-R-CA1"

/lab_host="DHI0B (Life Technologies)"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-CA1

library is a subtracted library derived from the following

tissues: thalamus, cerebellum, hypothalamus, medulla, pons

, midbrain, cerebral cortex, corpus striatum, testis, and

hippocampus. For a detailed description of the library

from which this clone was derived, please visit our web

site at ratest.eng.uiowa.edu. The subtraction has been

previously described in (Bonaldo, Lennon and Soares,

Genome Research 6:791-806, 1996)

TAG_LIB=UI-R-CA1

TAG_TISSUE=pons

TAG_SEQ=AGCACG

BASE COUNT 116 a 156 c 131 g 165 t

ORIGIN

Query Match 6.3%; Score 22; DB 12; Length 568;

Best Local Similarity 100.0%; Pred. No. 0.75;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 289

CAGTTCAGTGGAAACAGGACA 310

|||||

Db 550 CAGTTCAGTGGAAACAGGACA 529

RESULT 5

BB819780

LOCUS

BB819780 431 bp mRNA linear EST 19-NOV-2001

Jyg-MC(A) cDNA Mus musculus CDNA clone GB30002B05 3', mRNA

sequence.

ACCESSION

BB819780

VERSION

BB819780.1

KEYWORDS

EST.

SOURCE

ORGANISM

Mus musculus

house mouse.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 431)

REFERENCE

AUTHORS

Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T.,

Hayashizaki, T., Hirose, K., Hirose, T., Hirose, T., Imotani, K., Ishii

, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T.,

Nakamura, M., Nishikawa, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T.,

Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sakaki, D., Sato, K.,

Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagawa

, A., Takahashi, F., Takaku-Akai, S., Tanaka, T., Tomaru, A., Toya, T.,

Watahiki, A., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.

2001)

TITLE

JOURNAL

COMMENT

Unpublished (2001)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.

FEATURES source

```
1. .431
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="G830002B05"
/clone_lib="RIKEN full-length enriched, mammary gland
RCB-0526 Jyg-MC(A) cDNA"
/tissue_type="mammary gland"
/cell_line="RCB-0526 Jyg-MC(A)"
/notes="pooled cell lines; (cell_line=CRL-1751 WEHI 164),
(cell_line=CRL-2116 JC), (cell_line=RCB-0035 WEHI-3),
(cell_line=RCB-0464 Meth-A), (cell_line=RCB-0545 OHTA),
(cell_line=RCB-0559 K-1 F1), (cell_line=RCB-1283 B16
melanoma), (cell_type=B cells, cell_line=CRL-1702 WEHI 231
), (cell_type=Teydig cells, cell_line=CRL-2065 M12C-1),
(cell_type=Nullipotent stem cell, cell_line=CRL-2070 NE),
(tissue_type=bladder, cell_line=RCB-0544 MBT-2),
(tissue_type=bone marrow, cell_type=stroma cell,
cell_line=RCB-0549 Cle-H3), (tissue_type=kidney,
cell_line=CCL-142 RAG), (tissue_type=submandibular gland,
cell_line=CRL-1734 SCA-9 clone 15), (strain=BALB/C,
cell_type=B cells, cell_line=CRL-1669 BCL1 Clone 13.20-3B3
), (strain=C3H, tissue_type=brain, cell_line=CRL-1443
BC3H1)"
```

```
BASE COUNT 113 a 89 c 94 g 135 t
ORIGIN
```

```
Query Match 5.7%; Score 20; DB 10; Length 431;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 205 CTGGAGAAATTAGAGCCCC 224
|||||
Db 86 CTGGAGAAATTAGAGCCCC 105
```

```
RESULT 6
H58973/c 404 bp mRNA linear EST 06-OCT-1995
LOCUS
DEFINITION
IMAGE:207735 5', mRNA sequence.
ACCESSION H58973
VERSION H58973.1 GI:1011805
KEYWORDS EST.
SOURCE human.
```

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 404)
AUTHORS
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman
M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,
Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisan, E., Waterston
R., Williamson, A., Wohlmann, P. and Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
Insert Size: 864
High quality sequence stops: 294
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 864 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 294.
Location/Qualifiers
1. .404

FEATURES source

```
/organism="Homo sapiens"
/db_xref="GDB:3776866"
/db_xref="taxon:9606"
/clone="IMAGE:207735"
/clone_lib="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="D10B (ampicillin resistant)"
/notes="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5' AACTGGAAGATTAATAAGATCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
```

```
BASE COUNT 82 a 128 c 88 g 103 t
ORIGIN
```

```
Query Match 5.4%; Score 19; DB 14; Length 404;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 333 TGAGGCGCTGCTGGGCGCTG 351
|||||
Db 247 TGAGGCGCTGCTGGGCGCTG 229
```

```
RESULT 7
AI400688/c 416 bp mRNA linear EST 08-FEB-1999
LOCUS
DEFINITION
IMAGE:207735 x1 NCI_CGAP_CLL1 Homo sapiens cDNA clone IMAGE:2116344 3',
mRNA sequence.
ACCESSION AI400688
VERSION AI400688.1 GI:4243775
KEYWORDS EST.
SOURCE human.
```

```
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 416)
AUTHORS
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
```

Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever, M.D., Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html
 Seq primer: -40UP from Glibco
 High quality sequence stop: 342.

FEATURES

Location/Qualifiers

SOURCE

1. .416
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2116344"
 /clone_lib="NCI_CGAP_CLL1"
 /tissue_type="B-cell, chronic lymphocytic leukemia"
 /lab_host="DH10B"
 /note="Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand was primed with a Not I - oligo(dT) primer [5', TGTACCAATCTGAAGTGGAGCGCGCATTCCTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized, and was constructed by Bento Soares and M. Fatima Bonaldo."
 BASE COUNT 93 a 92 c 84 g 147 t
 ORIGIN

Query Match 5.4%; Score 19; DB 9; Length 416;

Best Local Similarity 100.0%; Pred. No. 30;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 286 GTGCAGTTCAGTGGAAC 304

|||||

Db 298 GTGCAGTTCAGTGGAAC 280

RESULT 8

AW178252/c

LOCUS AW178252 525 bp mRNA linear EST 16-NOV-1999
 DEFINITION MR0-HT0066-010999-001-e05 HT0066 Homo sapiens cDNA, mRNA sequence.
 ACCESSION AW178252
 VERSION AW178252.1 GI:6444289
 KEYWORDS EST.
 SOURCE human.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 525)
 HCGP http://www.ludwig.org.br/ORESTES.
 TITLE The FAPESP/LICR Human Cancer Genome Project
 JOURNAL Unpublished (1999)
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

REFERENCE

Ash Alizadeh, John Byrd, M.D., Mike Grever, M.D., Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html
 Seq primer: -40UP from Glibco
 High quality sequence stop: 342.

SOURCE

1. .539
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3078690"
 /clone_lib="NIH_MGC_50"
 /tissue_type="lymph"
 /cell_type="germinal center B cells"
 /cell_line="MGC85"
 /lab_host="DH10B (LTI)"
 /note="Vector: pT7T3-Pac; Site_1: NotI; Site_2: Eco RI; Constructed from size fractionated cytoplasmic mRNA (3.5-4.4Kb). Directionally cloned. Cells provided by Louis M. Staudt, Ph.D. Library preparation by Maria de Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."
 BASE COUNT 158 a 126 c 108 g 147 t
 ORIGIN

Query Match 5.4%; Score 19; DB 10; Length 539;

Best Local Similarity 100.0%; Pred. No. 33;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 286 GTGCAGTTCAGTGGAAC 304

|||||

FEATURES

Location/Qualifiers

SOURCE

1. .525
 /organism="Homo sapiens"

/db_xref="taxon:9606"
 /clone_lib="HT0066"
 /dev_stage="Adult"
 /note="Organ: head_neck; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 247 a 57 c 119 g 102 t
 ORIGIN

Query Match 5.4%; Score 19; DB 10; Length 525;

Best Local Similarity 100.0%; Pred. No. 33;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 TCCTGCTCCAGTTTCCA 118

|||||

Db 126 TCCTGCTCCAGTTTCCA 108

RESULT 9

AW503344

LOCUS AW503344 539 bp mRNA linear EST 01-MAR-2000
 DEFINITION UI-RF-BN0-akk-f-10-0-UI-r1 NIH_MGC_50 Homo sapiens cDNA clone
 IMAGE:3078690 5', mRNA sequence.
 ACCESSION AW503344
 VERSION AW503344.1 GI:7118645
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 539)
 NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Eco RI site shown at the beginning of the sequence.
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: M.B. Soares Lab
 cDNA Library Arrayed by: M.B. Soares Lab
 DNA Sequencing by: M.B. Soares Lab
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html
 Seq primer: M13 Forward

Location/Qualifiers

1. .539

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3078690"

/clone_lib="NIH_MGC_50"

/tissue_type="lymph"

/cell_type="germinal center B cells"

/cell_line="MGC85"

/lab_host="DH10B (LTI)"

/note="Vector: pT7T3-Pac; Site_1: NotI; Site_2: Eco RI; Constructed from size fractionated cytoplasmic mRNA (3.5-4.4Kb). Directionally cloned. Cells provided by Louis M. Staudt, Ph.D. Library preparation by Maria de Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

BASE COUNT 158 a 126 c 108 g 147 t

ORIGIN

Query Match 5.4%; Score 19; DB 10; Length 539;

Best Local Similarity 100.0%; Pred. No. 33;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 286 GTGCAGTTCAGTGGAAC 304

|||||

```
Db 412 GTGAGTTCACGTGGAAAC 430

RESULT 10
BE288244/c
LOCUS 741 bp mRNA linear EST 26-OCT-2000
DEFINITION 601095415f1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:348986 5',
mrna sequence.
ACCESSION BE288244
VERSION BE288244.1 GI:9167806
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 741)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM8531 row: n column: 03
High quality sequence start: 47
High quality sequence stop: 95.
Location/Qualifiers
1. .741
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:348986"
/clone_lib="NCI_CGAP_Mam5"
/tissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH"

BASE COUNT 142 a 214 c 189 g 193 t 3 others
ORIGIN

Query Match 5.4%; Score 19; DB 10; Length 741;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 286 GTGAGTTCACGTGGAAAC 304
|||||
Db 265 GTGAGTTCACGTGGAAAC 247

RESULT 11
BM051508
LOCUS 784 bp mRNA linear EST 07-NOV-2001
DEFINITION 603638192f1 NTH_MGC_8 Homo sapiens cDNA clone IMAGE:5419438 5',
mrna sequence.
ACCESSION BM051508
VERSION BM051508.1 GI:16780775
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 784)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)

Db 412 GTGAGTTCACGTGGAAAC 430

JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1872 row: g column: 23
High quality sequence stop: 748.
Location/Qualifiers
1. .784
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5419438"
/clone_lib="NIH_MGC_8"
/tissue_type="Burkitt lymphoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

```

/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: prostate; Vector: pDNR-LIB (Clontech);
Site_1: SfII (ggcgcctcgcc); Site_2: SfiI (ggccattatggcc
); 5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CAGGCCATTATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCGGCGGCACATG-dt(30)BN-3'
(where B = A, C, G, or T). Average
insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
BASE COUNT      334 a      89 c      224 g      198 t
ORIGIN
Query Match      5.4%; Score 19; DB 12; Length 845;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 TCTGTCTCTCCAGTTTCCA 118
|||||
Db 613 TCTGTCTCTCCAGTTTCCA 595
|||||

RESULT 13
BF680728/c
LOCUS
DEFINITION
602155347F1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4296395 5',
mRNA sequence.
ACCESSION
BF680728
VERSION
BF680728.1 GI:11954623
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 899)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cyapbs-r@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM148 row: n column: 12
High quality sequence stop: 615.
Location/Qualifiers
1. .899
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4296395"
/clone_lib="NIH_MGC_83"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: prostate; Vector: pDNR-LIB (Clontech);
Site_1: SfII (ggcgcctcgcc); Site_2: SfiI (ggccattatggcc
); 5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CAGGCCATTATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCGGCGGCACATG-dt(30)BN-3'
(where B = A, C, G, or T). Average
insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
BASE COUNT      327 a      139 c      236 g      197 t
ORIGIN
Query Match      5.4%; Score 19; DB 12; Length 899;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: prostate; Vector: pDNR-LIB (Clontech);
Site_1: SfII (ggcgcctcgcc); Site_2: SfiI (ggccattatggcc
); 5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CAGGCCATTATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCGGCGGCACATG-dt(30)BN-3'
(where B = A, C, G, or T). Average
insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
BASE COUNT      334 a      89 c      224 g      198 t
ORIGIN
Query Match      5.4%; Score 19; DB 12; Length 845;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 TCTGTCTCTCCAGTTTCCA 118
|||||
Db 613 TCTGTCTCTCCAGTTTCCA 595
|||||

RESULT 13
BF680728/c
LOCUS
DEFINITION
602155347F1 NIH_MGC_83 Homo sapiens cDNA clone IMAGE:4296395 5',
mRNA sequence.
ACCESSION
BF680728
VERSION
BF680728.1 GI:11954623
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 899)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cyapbs-r@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM148 row: n column: 12
High quality sequence stop: 615.
Location/Qualifiers
1. .899
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4296395"
/clone_lib="NIH_MGC_83"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: prostate; Vector: pDNR-LIB (Clontech);
Site_1: SfII (ggcgcctcgcc); Site_2: SfiI (ggccattatggcc
); 5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CAGGCCATTATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCGGCGGCACATG-dt(30)BN-3'
(where B = A, C, G, or T). Average
insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
BASE COUNT      327 a      139 c      236 g      197 t
ORIGIN
Query Match      5.4%; Score 19; DB 12; Length 899;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggcgcctcgcc); Site_2: SfiI (ggccattatggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CAGGCCATTATGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGGCGGCACATG-dt(30)BN-3' (where B = A,
C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."
BASE COUNT      363 a      182 c      270 g      192 t
ORIGIN
Query Match      5.4%; Score 19; DB 12; Length 1007;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 TCTGTCTCTCCAGTTTCCA 118
|||||
Db 640 TCTGTCTCTCCAGTTTCCA 622
|||||

RESULT 15
BM480049
LOCUS
DEFINITION
AGENCOURT_5424083 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5491592
5', mRNA sequence.
ACCESSION
BM480049
VERSION
BM480049.1 GI:18529091
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

QY 100 TCTGTCTCTCCAGTTTCCA 118
|||||
Db 633 TCTGTCTCTCCAGTTTCCA 615
|||||

RESULT 14
BG484905/c
LOCUS
DEFINITION
602503910F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4617170 5',
mRNA sequence.
ACCESSION
BG484905
VERSION
BG484905.1 GI:13417184
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1007)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cyapbs-r@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1373 row: d column: 03
High quality sequence stop: 650.
Location/Qualifiers
1. .1007
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4617170"
/clone_lib="NIH_MGC_77"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggcgcctcgcc); Site_2: SfiI (ggccattatggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CAGGCCATTATGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGGCGGCACATG-dt(30)BN-3' (where B = A,
C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."
BASE COUNT      363 a      182 c      270 g      192 t
ORIGIN
Query Match      5.4%; Score 19; DB 12; Length 1007;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 TCTGTCTCTCCAGTTTCCA 118
|||||
Db 640 TCTGTCTCTCCAGTTTCCA 622
|||||

RESULT 15
BM480049
LOCUS
DEFINITION
AGENCOURT_5424083 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5491592
5', mRNA sequence.
ACCESSION
BM480049
VERSION
BM480049.1 GI:18529091
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

REFERENCE
1 (bases 1 to 1575)
AUTHORS
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: rgs@bbs.fda.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM12112 row: f column: 09
High quality sequence stop: 372.

FEATURES
source
Location/Qualifiers
1..1575
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5491592"
/clone_lib="NIH_MGC_67"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."
BASE COUNT 434 a 533 c 270 g 333 t 5 others
ORIGIN

Query Match 5.4%; Score 19; DB 13; Length 1575;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 CACCATCATGCATGGCAAT 30
| | | | | | | | | | | | | | | |
Db 734 CACCATCATGCATGGCAAT 752

Search completed: November 6, 2002, 18:11:53
Job time : 2220 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 6, 2002, 17:40:37 ; Search time 89 Seconds
(without alignments)

1313.903 Million cell updates/sec

Title: US-09-897-438B-1

Perfect score: 351

Sequence: 1 gagcagtggtggaccatcat.....atgagcgctgctgggcccctg 351

Scoring table:

OLIGO_NUC
Gapop 60.0 , Capext 60.0

Searched: 310279 seqs, 166577418 residues

Word size : 0

Total number of hits satisfying chosen parameters: 620558

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications, NA.*

- 1: /cgn2.6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2.6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2.6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2.6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2.6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2.6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2.6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2.6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2.6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 10: /cgn2.6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2.6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 12: /cgn2.6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
- 13: /cgn2.6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 14: /cgn2.6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	351	100.0	351	10	US-09-897-438B-1
2	26	7.4	11580	10	US-09-880-107-3436
3	22	6.3	22	10	US-09-897-438B-6
4	21	6.0	21	10	US-09-897-438B-10
5	18	5.1	18	10	US-09-897-438B-11
6	18	5.1	198285	10	US-09-880-107-3814
7	17	4.8	325	10	US-09-764-877-2810
8	17	4.8	429	10	US-09-880-107-2470
9	17	4.8	499	10	US-09-864-761-14497
10	17	4.8	559	10	US-09-864-761-9458
11	17	4.8	675	10	US-09-764-864-678
12	17	4.8	1488	10	US-09-758-652-11
13	17	4.8	1732	10	US-09-923-556-1
14	17	4.8	1732	10	US-09-923-556-5
15	17	4.8	2131	10	US-09-764-864-250
16	17	4.8	3400	10	US-09-987-025-1
17	17	4.8	11881	10	US-09-764-869-2124
18	17	4.8	16552	10	US-09-764-855-321
19	17	4.8	16552	10	US-09-764-855-322

ALIGNMENTS

RESULT 1

US-09-897-438B-1

; Sequence 1, Application US/09897438B

; Patent No. US20020137095A1

; GENERAL INFORMATION:

; APPLICANT: Mikoshiba, Katsuhiko

; APPLICANT: Tate, Naoko

; TITLE OF INVENTION: REELIN PROTEIN CR-50 EPTOPE REGION

; FILE REFERENCE: 04853-0076-00000

; CURRENT APPLICATION NUMBER: US/09/897,438B

; CURRENT FILING DATE: 2001-07-03

; PRIOR APPLICATION NUMBER: JP 2000-202801

; PRIOR FILING DATE: 2000-07-04

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 351

; TYPE: DNA

; ORGANISM: Mus musculus

US-09-897-438B-1

Query Match 100.0%; Score 351; DB 10; Length 351;

Best Local Similarity 100.0%; Pred No. 2.4e-181;

Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCAGTGTGGCACCACATCATGCAATGCTGTCACTTCTGTGAGCGGTACGGCCCT 60

|||||

Db 1 GAGCAGTGTGGCACCACATCATGCAATGCTGTCACTTCTGTGAGCGGTACGGCCCT 60

QY 61 CGAGAGCTGACCAACACATGCCTGGAACACAAACAGACATCTCTCTCCAGTTTCCATT 120

|||||

Db 61 CGAGAGCTGACCAACACATGCCTGGAACACAAACAGACATCTCTCTCCAGTTTCCATT 120

QY 121 GGTCAGGATCATGTCGATTAGTTACTCTGACCCAGCATCTGTGTCTATACGCCAAG 180

|||||

Db 121 GGTCAGGATCATGTCGATTAGTTACTCTGACCCAGCATCTGTGTCTATACGCCAAG 180

QY 181 AACAAATACCCCTGATTGGATTTCAGCTGGAGAAAATAGAGCCCTTCCCAATGTGAGCACA 240

|||||

Db 181 AACAAATACCCCTGATTGGATTTCAGCTGGAGAAAATAGAGCCCTTCCCAATGTGAGCACA 240

QY 241 GTCATCCACATCTGTACCTCCCGGAGAAACCCAGAGAGCGGTGCGAGTTCCAGTGG 300

|||||

Sequence 1444, Ap

Sequence 1818, Ap

Sequence 1819, Ap

Sequence 15602, A

Sequence 782, App

Sequence 10721, A

Sequence 29, Appl

Sequence 1885, Ap

Sequence 1004, Ap

Sequence 1, Appl

Sequence 20, Appl

Sequence 2, Appl

Sequence 10, Appl

Sequence 11, Appl

Sequence 4, Appl

Sequence 1, Appl

Sequence 1, Appl

Sequence 1565, Ap

Sequence 3, Appl

Sequence 3, Appl

Sequence 119, App

Sequence 119, App

Sequence 119, App

Sequence 1848, Ap

Sequence 1182, Ap

Sequence 114, App

241 GTATCCACATCTGTACTCCCGAGGAAGSCAAGGGGAGAGCGTCAGTTCCAGTGG 300 Ddb

301 AAACAGGACACCTCGAGTGGTGATGAGGCCCTCTGGGCCCTG 351 QY

301 AAACAGACAGCTCGAGTGGGTGAGGTGTATAGGCCCTGTGGGCCCTG 351 Ddb

```

RESULT 2
US-09-880-107-3436
; Sequence 3436, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 1492-1-3068-00
; CURRENT APPLICATION NUMBER: US/09/880.107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3436
; LENGTH: 11580
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U75
US-09-880-107-3436

```

```

Query Match          7.4%; Score 26; DB 10; Length 11580;
Best Local Similarity 100.0%; Pred. No. 0.00014;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

RESULT 3
US-09-897-438B-6/c
; Sequence 6, Application US/09897438B
; Patent No. US20020137095A1
; GENERAL INFORMATION:
; APPLICANT: Mikoshiba, Katsuhiko
; APPLICANT: Tate, Naoko
; TITLE OF INVENTION: REELIN PROTEIN CR-50 EPTOPE REGION
; FILE REFERENCE: 04853-0076-00000
; CURRENT APPLICATION NUMBER: US/09/897,438B
; CURRENT FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: JP 2000-202801
; PRIOR FILING DATE: 2000-07-04
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: primer for PCR
US-09-897-438B-6

```

```

Db      22  GTATGAGCCTGCTGGGCCCTG 1
|||||
RESULT 4
US-09-897-438B-10
; Sequence 10, Application US/09897438B
; Patent No. US20020137095A1
; GENERAL INFORMATION:
; APPLICANT: Mikoshiba, Katsuhiko
; APPLICANT: Tate, Naoko
; TITLE OF INVENTION: RESLIN PROTEIN CR-50 EPIOTOPE REGION
; FILE REFERENCE: 04853-0076-00000
; CURRENT APPLICATION NUMBER: US/09/897,438B
; CURRENT FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: JP 2000-202801
; PRIOR FILING DATE: 2000-07-04
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: primer for PCR
US-09-897-438B-10

```

```

RESULT 5
US-09-897-438B-11
; Sequence 11, Application US/09897438B
; Patent No. US20020137095A1
; GENERAL INFORMATION:
; APPLICANT: Mikoshiba, Katsuhiko
; APPLICANT: Tate, Naoko
; TITLE OF INVENTION: REELIN PROTEIN CR-50 EPIOTOPE REGION
; FILE REFERENCE: 04853-0076-00000
; CURRENT APPLICATION NUMBER: US/09/897,438B
; CURRENT FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: JP 2000-202801
; PRIOR FILING DATE: 2000-07-04
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: primer for PCR
US-09-897-438B-11

```

RESULT 6
US-09-880-107-3814/c
; Sequence 3814, Application US/09880107

; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US-60/237,054
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3814
; LENGTH: 198285
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 X87344
US-09-880-107-3814

Query Match 5.1%; Score 18; DB 10; Length 198285;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 CTGACCACCATGCGCTG 84
|||||
Db 9594 CTGACCACCATGCGCTG 9577

RESULT 7
US-09-764-877-2810/c
; Sequence 2810, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2810
; LENGTH: 325
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-2810

Query Match 4.8%; Score 17; DB 10; Length 325;
Best Local Similarity 100.0%; Pred. No. 9.4;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 GCACCATCATGCGGC 27
|||||
Db 165 GCACCATCATGCGGC 149

RESULT 8
US-09-880-107-2470/c
; Sequence 2470, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107

; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2470
; LENGTH: 429
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 N27334
; NAME/KEY: unsure
; LOCATION: (1)..(429)
; OTHER INFORMATION: n = a or c or g or t
US-09-880-107-2470

Query Match 4.8%; Score 17; DB 10; Length 429;
Best Local Similarity 100.0%; Pred. No. 9.6;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 335 AGGCCTGCTGGCCCTG 351
|||||
Db 276 AGGCCTGCTGGCCCTG 260

RESULT 9
US-09-864-761-14497/c
; Sequence 14497, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21

; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 14497
; LENGTH: 499
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC000121.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.62
US-09-864-761-14497

Query Match 4.8%; Score 17; DB 10; Length 499;
Best Local Similarity 100.0%; Pred. No. 9.7; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 AGAGCCCTTCCAATGT 233
|||||
Db 487 AGAGCCCTTCCAATGT 471

RESULT 10
US-09-864-761-9458
; Sequence 9458, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203

; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 9458
; LENGTH: 559
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL022318.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.9
US-09-864-761-9458

Query Match 4.8%; Score 17; DB 10; Length 559;
Best Local Similarity 100.0%; Pred. No. 9.7; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 GAGGCTGTGGGCGCT 350
|||||
Db 208 GAGGCTGTGGGCGCT 224

RESULT 11
US-09-764-864-678/c
; Sequence 678, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 678
; LENGTH: 675
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-864-678

Query Match 4.8%; Score 17; DB 10; Length 675;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 TCTGCTCCAGTTTC 116
|||||
Db 107 TCTGCTCCAGTTTC 91

RESULT 12
US-09-758-652-11
; Sequence 11, Application US/09758652
; Patent No. US20010011377A1
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; GARY MICHAEL FADER
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; CLASSES OF SOYBEAN SEED
; PROTEIN GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA

ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS 95
SOFTWARE: MICROSOFT WORD 7.0A
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/758,652
FILING DATE: 11-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/019,940
FILING DATE: JUNE 14, 1996
ATTORNEY/AGENT INFORMATION:
NAME: LYNNE M. CHRISTENBURY
REGISTRATION NUMBER: 30,971
REFERENCE/DOCKET NUMBER: BB-1071-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-992-5481
TELEFAX: 302-773-0164
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1488 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-758-652-11

Query Match 4.8%; Score 17; DB 10; Length 1488;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 261 CCCGAGGAAGCCAAAG 277
|||||
DB 898 CCCGAGGAAGCCAAAG 914

RESULT 13

US-09-923-556-1/c
Sequence 1, Application US/09923556
Patent No. US20020058321A1
GENERAL INFORMATION:
APPLICANT: Busch, Marco
APPLICANT: Hain, Rudiger
APPLICANT: Martin, William
APPLICANT: Tietjen, Klaus
APPLICANT: Kloti, Andreas
TITLE OF INVENTION: Method of determining the activity of
TITLE OF INVENTION: 1-deoxy-D-xylose-5-phosphate reductoisomerase and
TITLE OF INVENTION: 1-deoxy-D-xylose-5-phosphate synthase
FILE REFERENCE: 2020US
CURRENT APPLICATION NUMBER: US/09/923,556
CURRENT FILING DATE: 2001-08-07
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/449,335
PRIOR FILING DATE: EARLIER FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1732
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: CDS
LOCATION: (18)..(1448)
OTHER INFORMATION: DNA encoding 1-deoxy-D-xylose-5-phosphate
OTHER INFORMATION: reductoisomerase
US-09-923-556-1

Query Match 4.8%; Score 17; DB 10; Length 1732;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 195 TTGGATTACGCTGGAGA 211
|||||
DB 55 TTGGATTACGCTGGAGA 39

RESULT 14

US-09-923-556-5/c
Sequence 5, Application US/09923556
Patent No. US20020058321A1
GENERAL INFORMATION:
APPLICANT: Busch, Marco
APPLICANT: Hain, Rudiger
APPLICANT: Martin, William
APPLICANT: Tietjen, Klaus
APPLICANT: Kloti, Andreas
TITLE OF INVENTION: Method of determining the activity of
TITLE OF INVENTION: 1-deoxy-D-xylose-5-phosphate reductoisomerase and
TITLE OF INVENTION: 1-deoxy-D-xylose-5-phosphate synthase
FILE REFERENCE: 2020US
CURRENT APPLICATION NUMBER: US/09/923,556
CURRENT FILING DATE: 2001-08-07
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/449,335
PRIOR FILING DATE: EARLIER FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 1732
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: CDS
LOCATION: (18)..(1448)
OTHER INFORMATION: new coding sequence
US-09-923-556-5

Query Match 4.8%; Score 17; DB 10; Length 1732;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 195 TTGGATTACGCTGGAGA 211
|||||
DB 55 TTGGATTACGCTGGAGA 39

RESULT 15

US-09-764-864-250/c
Sequence 250, Application US/09764864
Patent No. US20020132753A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT23
CURRENT APPLICATION NUMBER: US/09/764,864
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 1792
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 250
LENGTH: 2131
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-864-250

Query Match 4.8%; Score 17; DB 10; Length 2131;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 TCTGTCTCTCCAGTTTTC 116
|||||
DB 90 TCTGTCTCTCCAGTTTTC 74

Search completed: November 6, 2002, 19:26:17
Job time : 186 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 6, 2002, 15:28:31 ; Search time 2766 Seconds
(without alignments)
3693.089 Million cell updates/sec
Title: US-09-897-438b-1
Perfect score: 351
Sequence: 1 gagcagtggcaccatcat.....atgagcctgctggccctg 351
Scoring table: OLIGO_NUC
Gapox 60.0 , Gapext 60.0
Searched: 2054640 seqs, 14551402878 residues
Word size : 0
Total number of hits satisfying chosen parameters: 4109280
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Listing first 45 summaries

Database :

- GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	351	100.0	11673	6	AX305377	AX305377 Sequence
2	351	100.0	11673	10	MMU24703	U24703 Mus musculus
3	135	38.5	163985	10	AC121878	AC121878 Mus muscu
4	135	38.5	183156	2	AC023062	AC023062 Mus muscu
5	33	9.4	126130	2	AC095877	AC095877 Rattus no
6	33	9.4	183641	2	AC124933	AC124933 Rattus no
7	33	9.4	202764	2	AC128022	AC128022 Rattus no
8	32	9.1	11187	10	AB049473	AB049473 Rattus no
9	26	7.4	11580	6	AX410790	AX410790 Sequence
10	26	7.4	11580	9	HSU79716	U79716 Human reeli
11	26	7.4	11580	11	G30936	G30936 sws3176 Er
12	26	7.4	11580	11	G30938	G30938 sws3176 Er
13	24	6.8	70905	2	AC119906	AC119906 Mus muscu
14	21	6.0	73070	9	AL590143	AL590143 Human DNA
15	21	6.0	98873	2	AC015229	AC015229 Drosophil
16	21	6.0	160883	2	AL158814	AL158814 Homo sapi
17	21	6.0	168652	3	AC009384	AC009384 Drosophil
18	21	6.0	169059	2	AC009790	AC009790 Homo sapi
19	21	6.0	192491	2	AC103705	AC103705 Homo sapi
20	21	6.0	298386	3	AE003514	AE003514 Drosophil
21	20	5.7	6310	5	ACSWSP93	AF134194 Anolis ca
22	20	5.7	46890	2	AC109654	AC109654 Rattus no
23	20	5.7	128913	2	AP004053	AP004053 Oryza sat
24	20	5.7	164192	9	AC074338	AC074338 Human Chr
25	20	5.7	176163	2	AC099604	AC099604 Mus muscu
26	20	5.7	185652	9	AC007312	AC007312 Homo sapi
27	20	5.7	208405	2	AC079477	AC079477 Mus muscu
28	19	5.4	937	9	HSDBHGT	X63418 H.sapiens (
29	19	5.4	2796	9	AK095398	AK095398 Homo sapi
30	19	5.4	3452	9	AK091510	AK091510 Homo sapi
31	19	5.4	4883	6	BD012821	BD012821 Y568, a n
32	19	5.4	4883	23	BD008256	BD008256 Y568, a n
33	19	5.4	5092	9	HSMB04804	AL833491 Homo sapi
34	19	5.4	6931	9	HSMB03520	AL832213 Homo sapi
35	19	5.4	7215	6	BD012826	BD012826 Y568, a n
36	19	5.4	7215	6	BD012827	BD012827 Y568, a n
37	19	5.4	7215	9	AB059277	AB059277 Homo sapi
38	19	5.4	7215	23	BD008261	BD008261 Y568, a n
39	19	5.4	7215	23	BD008262	BD008262 Y568, a n
40	19	5.4	37584	9	AC001227	AC001227 Genomic s
41	19	5.4	44089	9	AC002101	AC002101 Genomic s
42	19	5.4	61676	2	AC060800	AC060800 Homo sapi
43	19	5.4	101555	2	AC095962	AC095962 Rattus no
44	19	5.4	144581	2	AC124868	AC124868 Rattus no
45	19	5.4	157266	2	AC127765	AC127765 Rattus no

ALIGNMENTS

RESULT 1	AX305377	11673 bp	DNA	linear	PAT 11-DEC-2001
LOCUS	Sequence 128 from Patent WO0188188.				
DEFINITION	AX305377				
ACCESSION	AX305377				
VERSION	AX305377.1	GI:17644926			
KEYWORDS					
SOURCE	house mouse.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
TITLE	Ishikawa, K., Asai, S., Takahashi, Y., Nagata, T. and Ishii, Y.				
JOURNAL	Method for examining ischemic conditions				
	Patent: WO 0188188-A 128 22-NOV-2001;				

Query Match 100.0%; Score 351; DB 10; Length 11673;
Best Local Similarity 100.0%; Pred. No. 9.7e-193;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCAGTGTGGCACCATCATGCGAATGCTGTACCTTCTGTGAGCGGTAGCGCCCT 60
|||||
Db 970 GAGCAGTGTGGCACCATCATGCGAATGCTGTACCTTCTGTGAGCGGTAGCGCCCT 1029
|||||

QY 61 CGAGAGTGCACCAACCATCCCTGAACACAAACACAGCATCTGTCCPCCAGTTTCCATT 120
|||||
Db 1030 CGAGAGTGCACCAACCATCCCTGAACACAAACACAGCATCTGTCCPCCAGTTTCCATT 1089
|||||

QY 121 GGTCAGGATCATGTCGATTTAGTTACTCTGACCCAGCATCTGTGTATATAGGCCAAG 180
|||||
Db 1090 GGTCAGGATCATGTCGATTTAGTTACTCTGACCCAGCATCTGTGTATATAGGCCAAG 1149
|||||

QY 181 AACAAATACCCGCTGATTCGATTCAGCTGGAGAAATAGAGCCCTTCCAAATGTGAGCACA 240
|||||
Db 1150 AACAAATACCCGCTGATTCGATTCAGCTGGAGAAATAGAGCCCTTCCAAATGTGAGCACA 1209
|||||

QY 241 GTCATCCACATCTGTACCTCCCGAGGAGCAACCAAGGAGAGCGTGCAGTTCAGCTGG 300
|||||
Db 1210 GTCATCCACATCTGTACCTCCCGAGGAGCAACCAAGGAGAGCGTGCAGTTCAGCTGG 1269
|||||

QY 301 AAACAGACAGCCCTGCAGTGGGTGAGGTGTATGAGCCCTGCTGGGCCCTG 351
|||||
Db 1270 AAACAGACAGCCCTGCAGTGGGTGAGGTGTATGAGCCCTGCTGGGCCCTG 1320
|||||

RESULT 3
AC121878/c
LOCUS AC121878 163985 bp DNA linear ROD 21-JUN-2002
DEFINITION Mus musculus clone RP24-131C8, complete sequence.
ACCESSION AC121878
VERSION AC121878.1 GI:21039993
KEYWORDS HTG.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
AUTHORS 1 (bases 1 to 163985)
TITLE The sequence of Mus musculus clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 163985)
AUTHORS McPherson, J.D. and Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (21-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 3 (bases 1 to 163985)
AUTHORS McPherson, J.D. and Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (21-JUN-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
Contact: submissions@wustl.edu
----- Project Information -----
Center project name: M_BB0131C08
----- Location/Qualifiers -----
1. .163985
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP24-131C8"

BASE COUNT 50392 a 33960 c 32422 g 4721 t
ORIGIN

Query Match 38.5%; Score 135; DB 10; Length 163985;

Best Local Similarity 100.0%; Pred. No. 6e-67;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 AGAGCCCTTTCCAAATGTGAGCAGATCATCCACATCTGTACCTCCCGAGAGCCAAA 276
|||||
Db 15382 AGAGCCCTTTCCAAATGTGAGCAGATCATCCACATCTGTACCTCCCGAGAGCCAAA 15323
|||||

QY 277 GGGGAGAGCGTGCAGTTCCAGTGGAAACAGACAGCCCTGCAGTGGGTGAGGTATGAG 336
|||||
Db 15322 GGGGAGAGCGTGCAGTTCCAGTGGAAACAGACAGCCCTGCAGTGGGTGAGGTATGAG 15263
|||||

QY 337 GCCTGCTGGGCCCTG 351
|||||
Db 15262 GCCTGCTGGGCCCTG 15248
|||||

RESULT 4
AC023062 183156 bp DNA linear HTG 05-MAY-2000
LOCUS AC023062
DEFINITION Mus musculus chromosome 5 clone RP23-466D2 strain C57BL6/J, WORKING
DRAFT SEQUENCE, 8 unordered pieces.
ACCESSION AC023062
VERSION AC023062.2 GI:7709916
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Mus musculus.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W.,
Buffard, G.G., Dietrich, N.L., Eagle, W.O., Gupta, J., Ho, S.-L.,
Huang, M.C., Idol, J., Lee-Lin, S.-Q., Maduro, Q.L., Maduro, V.B.,
Mastriano, S.D., McCloskey, J.C., Morse, E., Ojodu, M.A., Pearson, R.,
Stantrip, S., Summers, T.J., Thomas, J.W., Thomas, P.J.,
Tongson, E.E., Touchman, J.W., Tran, J.T., Vogt, J.L., Walker, M.A.,
Wetherby, K.D. and Green, E.D.
TITLE NISC Mouse Sequencing Initiative
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 183156)
AUTHORS Green, E.D.
TITLE Direct Submission
JOURNAL Submitted (08-FEB-2000) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Galthersburg, MD 20877, USA
COMMENT On May 5, 2000 this sequence version replaced gi:6939141.
----- Genome Center -----
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: <http://www.nisc.nih.gov>
Contact: nisc_mouse@nhgri.nih.gov
----- Project Information -----
Center project name: Y1
Center clone name: 466D02
----- Summary Statistics -----
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 175371 bases at least Q40
Consensus quality: 178474 bases at least Q30
Consensus quality: 180133 bases at least Q20
Insert size: 179000; agarose-fp
Insert size: 188000; pulse-field-gel
Quality coverage: 6.96x in Q20 bases; agarose-fp
Quality coverage: 6.63x in Q20 bases; pulse-field-gel
Quality coverage: 6.80x in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will

```

* be preserved.
1 2374: contig of 2374 bp in length
* gap of unknown length
* 2375 4490: contig of 2116 bp in length
* gap of unknown length
* 4491 14959: contig of 10469 bp in length
* gap of unknown length
* 14960 31072: contig of 16113 bp in length
* gap of unknown length
* 31073 50119: contig of 19047 bp in length
* gap of unknown length
* 50120 78196: contig of 28077 bp in length
* gap of unknown length
* 78197 109874: contig of 31678 bp in length
* gap of unknown length
* 109875 183156: contig of 73282 bp in length.
FEATURES
    Location/Qualifiers
1..183156
    /organism="Mus musculus"
    /strain="C57BL6/J"
    /db_xref="taxon:10090"
    /chromosome="5"
    /clone="RP23-466D2"
    /clone_lib="RPCI mouse BAC library 23"
BASE COUNT 53719 a 37080 c 37302 g 55052 t 3 others
ORIGIN

Query Match      38.5%; Score 135; DB 2: Length 183156;
Best Local Similarity 100.0%; Pred. No. 6e-67;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 AGAGCCCTTCCAAATGTGAGCACAGTATCCACATCTCTGTACCTCCCGAGGAGCCAAA 276
      |||||
Db 173623 AGAGCCCTTCCAAATGTGAGCACAGTATCCACATCTCTGTACCTCCCGAGGAGCCAAA 173682

QY 277 GGGGAGAGCGTGACATCCAGTGAACAGCAGCCCTGGAGTGGCTGAGGTGTATGAG 336
      |||||
Db 173683 GGGGAGAGCGTGACATCCAGTGAACAGCAGCCCTGGAGTGGCTGAGGTGTATGAG 173742

QY 337 GCCTGCTGGGCCCTG 351
      |||||
Db 173743 GCCTGCTGGGCCCTG 173757

RESULT 5
AC095877
LOCUS      126130 bp DNA linear HTG 11-JUL-2002
DEFINITION Rattus norvegicus clone CH230-10G13, *** SEQUENCING IN PROGRESS
            ***, 59 unordered pieces.
AC095877
VERSION    AC095877.4 GI:21722955
KEYWORDS   HTG; HTGS_PHASE1.
SOURCE     Norway rat.
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
            1 (bases 1 to 126130)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Alisbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,

```

```

Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlisson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,M., Nguyen,S.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okwionu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I.,
Sodergren,E., Sonaikhe,I., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 126130)
Worley,K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 126130)
Worley,K.C.
Direct Submission
Submitted (11-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 10, 2002 this sequence version replaced gi:20975941.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GDRS
Center clone name: CH230-10G13
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 70934 bases at least Q40
Consensus quality: 75985 bases at least Q30
Consensus quality: 80731 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 59 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1382: contig of 1382 bp in length
* 1383 1482: gap of unknown length
* 1483 3046: contig of 1564 bp in length
* 3047 3146: gap of unknown length
* 3147 4489: contig of 1343 bp in length
* 4490 4589: gap of unknown length
* 4590 5786: contig of 1197 bp in length

```



```
* 5787 5886: gap of unknown length
* 5887 7115: contig of 1229 bp in length
* 7116 7115: gap of unknown length
* 7216 8771: contig of 1556 bp in length
* 8772 8771: gap of unknown length
* 8872 10260: contig of 1389 bp in length
* 10261 10360: gap of unknown length
* 10361 11370: contig of 1010 bp in length
* 11371 11470: gap of unknown length
* 11471 12710: contig of 1240 bp in length
* 12711 12810: gap of unknown length
* 12811 14180: contig of 1370 bp in length
* 14181 14280: gap of unknown length
* 14281 16048: contig of 1768 bp in length
* 16049 16148: gap of unknown length
* 16149 17569: contig of 1421 bp in length
* 17570 17659: gap of unknown length
* 17670 18820: contig of 1151 bp in length
* 18821 18920: gap of unknown length
* 18921 20209: contig of 1289 bp in length
* 20210 20309: gap of unknown length
* 20310 21633: contig of 1324 bp in length
* 21634 21733: gap of unknown length
* 21734 23269: contig of 1536 bp in length
* 23270 23369: gap of unknown length
* 23370 24647: contig of 1278 bp in length
* 24648 24747: gap of unknown length
* 24748 25800: contig of 1053 bp in length
* 25801 27017: contig of 1117 bp in length
* 27018 27117: gap of unknown length
* 27118 29183: contig of 2066 bp in length
* 29184 29283: gap of unknown length
* 29284 30315: contig of 1032 bp in length
* 30316 32086: contig of 1671 bp in length
* 32087 32186: gap of unknown length
* 32187 33395: contig of 1209 bp in length
* 33396 33495: gap of unknown length
* 33496 35663: contig of 2168 bp in length
* 35664 35763: gap of unknown length
* 35764 38044: contig of 2281 bp in length
* 38045 38144: gap of unknown length
* 38145 39845: contig of 1701 bp in length
* 39846 39945: gap of unknown length
* 39946 41982: contig of 2037 bp in length
* 41983 42083: gap of unknown length
* 42084 43743: contig of 1661 bp in length
* 43744 45904: contig of 2061 bp in length
* 45905 46004: gap of unknown length
* 46005 47667: contig of 1663 bp in length
* 47668 47767: gap of unknown length
* 47769 49833: contig of 2066 bp in length
* 49834 49933: gap of unknown length
* 49934 51434: contig of 1501 bp in length
* 51435 51534: gap of unknown length
* 51535 53778: contig of 2244 bp in length
* 53779 53878: gap of unknown length
* 53879 55552: contig of 1674 bp in length
* 55553 55652: gap of unknown length
* 55653 57670: contig of 2018 bp in length
* 57671 57770: gap of unknown length
* 57771 59630: contig of 1860 bp in length
* 59631 59730: gap of unknown length
* 59731 61147: contig of 1417 bp in length
* 61148 61247: gap of unknown length
* 61248 63995: contig of 2748 bp in length
* 63996 64095: gap of unknown length
* 64096 66294: contig of 2199 bp in length
* 66295 66394: gap of unknown length
* 66395 68794: contig of 2400 bp in length
* 68795 68894: gap of unknown length
```

```
* 68895 71533: contig of 2639 bp in length
* 71534 71633: gap of unknown length
* 71634 73469: contig of 1836 bp in length
* 73470 73569: gap of unknown length
* 73570 77054: contig of 3485 bp in length
* 77055 77154: gap of unknown length
* 77155 79005: contig of 1851 bp in length
* 79006 79105: gap of unknown length
* 79106 81294: contig of 2189 bp in length
* 81295 83864: gap of unknown length
* 83865 83964: contig of 2470 bp in length
* 83965 86039: contig of 2075 bp in length
* 86040 86139: gap of unknown length
* 86140 88646: contig of 2507 bp in length
* 88647 91823: contig of 3077 bp in length
* 91824 91923: gap of unknown length
* 91924 94369: contig of 2446 bp in length
* 94370 94469: gap of unknown length
* 94470 96886: contig of 2417 bp in length
* 96887 96986: gap of unknown length
* 96987 100679: contig of 3693 bp in length
* 100680 100779: gap of unknown length
* 100780 104171: contig of 3392 bp in length
* 104172 104271: gap of unknown length
* 104272 106760: contig of 2489 bp in length
```

Query Match 9.4%; Score 33; DB 2: Length 126130;

Best Local Similarity 100.0%; Pred. No. 1.5e-07; Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 285 CGTGCAGTTCCTCCAGTGGAAACAGACGCTGCG 317
      |||
Db 92895 CGTGCAGTTCCTCCAGTGGAAACAGACGCTGCG 92927
```

RESULT 6

AC124933/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AC124933 183641 bp DNA linear HTG 31-JUL-2002
Rattus norvegicus clone CH230-466N17, *** SEQUENCING IN PROGRESS
***, 88 unordered pieces.
AC124933.2 GI:21952716
HTG: HTGS_PHASE1.
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 183641)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Catroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabis,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homsí,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlssoon,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,

```

Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokwenkwo,S., Oguh,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansley,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wlecczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G., and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 183641)
Worley,K.C.
Direct Submission
Submitted (20-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 183641)
Worley,K.C.
Direct Submission
Submitted (31-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 24, 2002 this sequence version replaced gi:21490053.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KAVI
Center clone name: CH230-466N17
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 111724 bases at least Q40
Consensus quality: 120970 bases at least Q30
Consensus quality: 128458 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 88 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1107: contig of 1107 bp in length
* 1108: gap of unknown length
* 1208: contig of 1253 bp in length
* 2461: 2560: gap of unknown length
* 2561: 4004: contig of 1444 bp in length
* 4005: 4104: gap of unknown length
* 4105: 5382: contig of 1278 bp in length
* 5383: 5482: gap of unknown length
* 5483: 6874: contig of 1392 bp in length
* 6875: 6974: gap of unknown length
* 6975: 8071: contig of 1097 bp in length
* 8072: 8171: gap of unknown length
* 8172: 9478: contig of 1307 bp in length
* 9479: 9578: gap of unknown length
*
11061: contig of 1483 bp in length
11161: gap of unknown length
12742: contig of 1581 bp in length
12842: gap of unknown length
14504: contig of 1662 bp in length
14604: gap of unknown length
15734: contig of 1130 bp in length
15834: gap of unknown length
16958: contig of 1124 bp in length
17059: gap of unknown length
18704: contig of 1646 bp in length
18804: gap of unknown length
20420: contig of 1616 bp in length
20520: gap of unknown length
21648: contig of 1128 bp in length
21748: gap of unknown length
23177: contig of 1429 bp in length
23277: gap of unknown length
24760: contig of 1483 bp in length
24860: gap of unknown length
26759: contig of 1899 bp in length
26859: gap of unknown length
28239: contig of 1440 bp in length
28399: gap of unknown length
29592: contig of 1193 bp in length
29692: gap of unknown length
30870: contig of 1178 bp in length
30970: gap of unknown length
32642: contig of 1672 bp in length
32742: gap of unknown length
34386: contig of 1644 bp in length
34486: gap of unknown length
36064: contig of 1578 bp in length
36164: gap of unknown length
37398: contig of 1234 bp in length
37498: gap of unknown length
39608: contig of 2110 bp in length
39708: gap of unknown length
41480: contig of 1772 bp in length
41580: gap of unknown length
42759: contig of 1179 bp in length
42859: gap of unknown length
44711: contig of 1852 bp in length
44811: gap of unknown length
46208: contig of 1397 bp in length
46308: gap of unknown length
47767: contig of 1459 bp in length
47867: gap of unknown length
49227: contig of 1360 bp in length
49327: gap of unknown length
50797: contig of 1470 bp in length
50897: gap of unknown length
51928: contig of 1031 bp in length
52028: gap of unknown length
53105: contig of 1077 bp in length
53205: gap of unknown length
54546: contig of 1341 bp in length
54646: gap of unknown length
55820: contig of 1174 bp in length
55920: gap of unknown length
57635: contig of 1715 bp in length
57735: gap of unknown length
58942: contig of 1207 bp in length
59042: gap of unknown length
60341: contig of 1299 bp in length
60441: gap of unknown length
61693: contig of 1252 bp in length
61793: gap of unknown length
63315: contig of 1522 bp in length
63415: gap of unknown length
64571: contig of 1156 bp in length
64671: gap of unknown length
66083: contig of 1412 bp in length

```

```

* 66084 66183: gap of unknown length
* 66184 67516: contig of 1333 bp in length
* 67517 67516: gap of unknown length
* 67617 69820: contig of 2204 bp in length
* 69821 69820: gap of unknown length
* 69921 71261: contig of 1341 bp in length
* 71262 71361: gap of unknown length
* 71362 72775: contig of 1414 bp in length
* 72776 72875: gap of unknown length
* 72876 75298: contig of 2423 bp in length
* 75299 75398: gap of unknown length
* 75399 77094: contig of 1696 bp in length
* 77095 77194: gap of unknown length
* 77195 78334: contig of 1640 bp in length
* 78335 78934: gap of unknown length
* 78935 79989: contig of 1055 bp in length
* 79990 80089: gap of unknown length
* 80090 81774: contig of 1685 bp in length
* 81775 81874: gap of unknown length
* 81875 83255: contig of 1381 bp in length

Query Match          9.4%; Score 33; DB 2; Length 183641;
Best Local Similarity 100.0%; Pred. NO. 1.5e-07;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 285 CGTGCAGTTCAGTGAACAGACGACGCTGCG 317
Db 143654 CGTGCAGTTCAGTGAACAGACGACGCTGCG 143622
|||||
CGTGCAGTTCAGTGAACAGACGACGCTGCG 317
CGTGCAGTTCAGTGAACAGACGACGCTGCG 143622

RESULT 7
AC128022/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbata,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowles,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carton,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
Homsif,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jollivet,S., Joudah,S.,
Karlisson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newtonson,J., Newtonson,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
Oraquanye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I.,
Sodergren,E., Sonake,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabot,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczkyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 202764)
Worley,K.C.
Direct Submission
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KAIV
Center clone name: CH230-525L20
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 113069 bases at least Q40
Consensus quality: 119970 bases at least Q30
Consensus quality: 125656 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 97 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1244: contig of 1244 bp in length
* 1245 1344: gap of unknown length
* 1345 2346: contig of 1002 bp in length
* 2347 2446: gap of unknown length
* 2447 3813: contig of 1367 bp in length
* 3814 3913: gap of unknown length
* 3914 5101: contig of 1188 bp in length
* 5102 5201: gap of unknown length
* 5202 6724: contig of 1523 bp in length
* 6725 6824: gap of unknown length
* 6825 8048: contig of 1224 bp in length
* 8049 8148: gap of unknown length
* 8149 9212: contig of 1064 bp in length
* 9213 9312: gap of unknown length
* 9313 10432: contig of 1120 bp in length
* 10433 10532: gap of unknown length
* 10533 11890: contig of 1358 bp in length
* 11891 11990: gap of unknown length
* 11991 13068: contig of 1078 bp in length
* 13069 13168: gap of unknown length
* 13169 14486: contig of 1318 bp in length
* 14487 14587: gap of unknown length
* 14588 15996: contig of 1410 bp in length
* 15997 16096: gap of unknown length
* 16097 17534: contig of 1438 bp in length
* 17535 17634: gap of unknown length
* 17635 18703: contig of 1069 bp in length
* 18704 18803: gap of unknown length

```

* 18804 19884: contig of 1081 bp in length
* 19885 19885: gap of unknown length
* 19885 21539: contig of 1555 bp in length
* 21540 21639: gap of unknown length
* 21640 22780: contig of 1141 bp in length
* 22781 22880: gap of unknown length
* 22881 23926: contig of 1046 bp in length
* 23927 24026: gap of unknown length
* 24027 25544: contig of 1518 bp in length
* 25545 25644: gap of unknown length
* 25645 26899: contig of 1255 bp in length
* 26900 26999: gap of unknown length
* 27000 28237: contig of 1238 bp in length
* 28238 28337: gap of unknown length
* 28338 29758: contig of 1421 bp in length
* 29759 29858: gap of unknown length
* 29859 31038: contig of 1180 bp in length
* 31039 31138: gap of unknown length
* 31139 32279: contig of 1041 bp in length
* 32180 32279: gap of unknown length
* 32280 33796: contig of 1517 bp in length
* 33797 33896: gap of unknown length
* 33897 35380: contig of 1484 bp in length
* 35381 35480: gap of unknown length
* 35481 37032: contig of 1552 bp in length
* 37033 37132: gap of unknown length
* 37133 38414: contig of 1282 bp in length
* 38415 38514: gap of unknown length
* 38515 39948: contig of 1434 bp in length
* 39949 40048: gap of unknown length
* 40049 41436: contig of 1388 bp in length
* 41437 41536: gap of unknown length
* 41537 43085: contig of 1548 bp in length
* 43085 43184: gap of unknown length
* 43185 44940: contig of 1756 bp in length
* 44941 45040: gap of unknown length
* 45041 46186: contig of 1146 bp in length
* 46187 46286: gap of unknown length
* 46287 47989: contig of 1703 bp in length
* 47990 48089: gap of unknown length
* 48090 49842: contig of 1753 bp in length
* 49843 49942: gap of unknown length
* 49943 51128: contig of 1186 bp in length
* 51129 51228: gap of unknown length
* 51229 52823: contig of 1595 bp in length
* 52824 52924: gap of unknown length
* 52924 54431: contig of 1508 bp in length
* 54432 54532: gap of unknown length
* 54532 56042: contig of 1511 bp in length
* 56043 56142: gap of unknown length
* 56143 57287: contig of 1145 bp in length
* 57288 57387: gap of unknown length
* 57388 58993: contig of 1506 bp in length
* 58994 58993: gap of unknown length
* 58994 60024: contig of 1031 bp in length
* 60025 60124: gap of unknown length
* 60125 62858: contig of 2734 bp in length
* 62859 64365: contig of 1407 bp in length
* 64366 64465: gap of unknown length
* 64466 66434: contig of 1969 bp in length
* 66435 66534: gap of unknown length
* 66535 68632: contig of 2097 bp in length
* 68632 68731: gap of unknown length
* 68732 70839: contig of 2108 bp in length
* 70840 70939: gap of unknown length
* 70940 73009: contig of 2070 bp in length
* 73010 73109: gap of unknown length
* 73110 74557: contig of 1848 bp in length
* 74558 74958: gap of unknown length
* 74958 76816: contig of 1759 bp in length
* 76817 76916: gap of unknown length
* 76917 78679: contig of 1763 bp in length

* 78680 78779: gap of unknown length
* 78780 80056: contig of 1277 bp in length
* 80057 80156: gap of unknown length
* 80157 81312: contig of 1156 bp in length
* 81313 81412: gap of unknown length
* 81413 82922: contig of 1510 bp in length
* 82923 83022: gap of unknown length
* 83023 84180: contig of 1158 bp in length
* 84181 84280: gap of unknown length
* 84281 85849: contig of 1569 bp in length
* 85850 85949: gap of unknown length
* 85950 87612: contig of 1663 bp in length
* 87613 87712: gap of unknown length

Query Match 9.4%; Score 33; DB 2; Length 202764;
Best Local Similarity 100.0%; Pred.No. 1.5e-07;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 285 CGTGCAGTTCCAGTGGAAACAGGACAGCCTGCG 317
|||||
Db 187247 CGTGCAGTTCCAGTGGAAACAGGACAGCCTGCG 187215

RESULT 8
LOCUS AB049473
DEFINITION Rattus norvegicus mRNA for reelin, complete cds.
ACCESSION AB049473
VERSION AB049473.1 GI:17221617
KEYWORDS
SOURCE Rattus norvegicus juvenile cerebellum cDNA to mRNA.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1
Kikkawa,S. and Terashima,T.
rat reelin (Rein) complete CDS
Published Only in Database (2001)
REFERENCE 2 (bases 1 to 11187)
AUTHORS Kikkawa,S. and Terashima,T.
TITLE Direct Submission
JOURNAL Submitted (29-SEP-2000) Satoshi Kikkawa, Kobe University School of
Medicine, Department of Anatomy; Chuo-ku, Kusunoki-cho 7-5-1, Kobe,
Hyogo 650-0017, Japan (E-mail:skikkawa@med.kobe-u.ac.jp,
Tel:81-78-382-5325; Fax:81-78-382-5328)

FEATURES
source location/Qualifiers
1..11187
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/tissue_type="cerebellum"
/dev_stage="juvenile"
1..11187
/gene="reln"
341..10729
/gene="Rein"
/codon_start=1
/product="reelin"
/protein_id="BAB78470.1"
/db_xref="GI:17221618"
/translation="MERGWAPRTLVLAIVLALLLTLARAAATGYPRFSPFFFLCTH
HCELEGEQGEVLISLHAGNYTYVQEHVHTISTFFDGLLVLTGLTSTIQS
SQSIGGSFAFGFIMSDHQFNQFMCVAVSHLPTNLSFVFIAPAGTCGVNFM
ATATHRGQVIFKDALAQQLCEQAGTEATAYSHLAEIHSDSVILRRDDSDVHQLBLN
NIWBCSNDTGEQCGTLMHGNVTECEPYGPRLTTLNTTASVLFQSTGSGSCR
FSYSDPSLIIVYAKNNTADWLEKIRAPSNVSTIIHLYLPEDAKGENVQFQWRQDS
LHVGEYEAQWALDNILVNSAHQVILEDSLPDVTGNLFFPGATVKHSCQSDGNA
IYFHGNEGSQLNFATRDVLDSTIEQWSEEFESQPTGWDILGAVVSGCQTIESG
LSLVFLKDGKRGKLTGYGNLRFYFAMGCTCDPGDSDHENDVILYAKTEGKKEH
IALDTLSYSSYKVPPLVSVINPELQATKFCCLKQKHQGNHQNQNVAVDFHFLPIL
PSTSHMIQFINSILGCGTHQPGNSVLSFESINHGHSWLSLITCELCPCAGPHLPHST
IYSENYSGNWRVTIPLNAAALTRDTRIRWRQTGPIGNMMAIDNVYIGPSCLEKFCSG
RGQCTRHGCKDPCGFSFACEMASQTFFPMFISESGSSRLSSYHNHYSIRGAEVSGFC

GVLASGKALVFNKGRBLITSLFDSSQSRRFLOFTLRGSKSVLSTCRAPDQGEVCL
LHYSYDNGTIVKLLLEHYSYLNTHPRIISVELPDDAKQFGIOFRWQPIHSTSGQEDVW
ADEILMTSVLNSISLDTNLEVTYSLGFIYLNIOFYCGHDITLCFTGSKLASSM
RYVETQSGMIGASYMTQFSLVAGCGQKTYPHMDNQVKLEYSTNHGLTWHLVDECPLPS
MPSCQFTSASIIYHASEFTQWRVTVILPKQTSVGATFRFWSQSYTTAOWEALDIIY
IGQOCNPMCSGHGSCDHGCRDQGYOGTECHPEAALPSTIMSDRENPSWSDWQEV
IGGEVVKPQGGVGVSSGSLYFSKAGKRLYSWDLTSMWDEFVQFYIOIGESACN
KPDREGVLLQYSNNGGIQWILLAEYFSDRGKRPVYIELEPAAAKTCTCTFRFWAQ
FSGEDYDQAWADDIILISEKQVPIPVNPTLPONFYEKPAFPDIPINGMSVWLMIAN
EGMAKNDSTPSAMVEFGKSGDRFAVTRDLTLKPGYVLOFKLNGISQFSSSTAP
VLIOYSHDAGSMWFLVKEGCFAPASGCGEGNSRELSEPTVYITGTFEWRVTVTAIP
RLASKTRFRIOESSOKNVPFGLDGVYISEPCVSCSHGDCISCGCDLGYT
AASCTPLNINPHESEMFDRGKLSPLWKISGGOVGTCGCTLSGDRSLYFNGLRE
ARTVPLDTRNIRLVQFYIOIGKTSIGTICIKPRARNEGVLVQYSDNGLIHLHLLRELD
FLSFLPQIISIDLPREAKTAPAFRWWQPHQKHSQAQWALDDVLIGVNDSSQTFQD
KFDGSDIDQANWYRIGGQVDIDCLSDMTALIFTENIGKPRYAETWDFHVSASSTQF
DMSGSCKPFDSATHQYSLNNGKDRHPTVECPYITGCVHYHTESSTYTSERFQ
WRVTVYLPATNSPRTRFRWQIYQANVTMGADAWAIDNVLASGCPWLCSGRGICDSGR
CVCDRGFGPGFCVPVPLPSILKDDFNGNLHPDLWPEVYGAERNGLNGETIKSGTSLI
FKGEGRLMLISRLDCTNMYVQFSLRFTAKGTPERSHSILLOSSINGVYWRLMDEF
YFPQTSIILFVNPVLSAQTNATRLRQWPNYNGKKEIWIIDDFILOGDNLNNPVM
LLDTDFGPREDNWFFYPGNGIYGLYCPYSKGAPEEDSAMVFNSEIGHSITTRDLS
VENTIIIOFEINVGCTSDSSADPVRLEFRDFGATWHLLPLCYHSSLVSLCSTLS
HHPSTIYAGTTCQWRREVHFHKLHLCGSFRFYQGFYPAGSQPVTWADNVYIGP
QCEEMCGGSCVNGTKICDPYSGPTCKISTKNPDLKDDFEGOLESDRFLMSSG
KPSKCGIILSGNNLFENEDGLRMLVTRDLDLSHAREVQFMRGLGCGKGPDPERSQPY
LQOYSLNGLSLSLQEFILPSSNNGVRYTIALEMPLKARSQSTRLRWQPSENGHFYS
PWYDILIGLISGNVLEDDFSTLDSRKLWLLHPGGTKMPVCGSGDGLAVFTFEKAST
RVVYTDIAVNDESLQIDFAASCYSTDSYAIETLEYSDVLGSLHPLVRDCLPTNVE
CGRYHLQRLIVLSTFNKTRITPLPAYTRSOATRFRWHPAPEDKQQTWADNVYIG
DCGLDMCSGHGRICQSCVCDEQWGLCYDEPETSILPTOLKDNFNAPSNOWMLTVNG
GLSTVCGAVASGLAHFSGGCSRLLYTVDLNLTAETFOFYFMVGLCLITPSNRNGV
LLEYSVNGGITWLLMEIFYDQSKPGFVNILLPPDAKEIGTRFRWQPRHGDLDQND
WADNVLLISGSADQRTVMLDTFSAPVQHERSPADAPGVRIAFDMFEDRITSVNE
VYFHDDCTVERFCDSPGVLACSHDREVYAVTHDLTPTENNIMQIKYSVGCKVPEK
IAQNOIHVQSTDFGWSYVLPQCLPADPKCGTVSQSPVSFPFPTKGRITVPLPES
LMGNVRFRFYQKYSDVOWAIDNVFLYGLPCLDNCGHGDCKEKCICDPYSGPHCYL
THTLTKFLERDSEIKPDLWMSLEGGSTCEGLIAENTALIFYGGSVTRQAITQDL
DURKAFIDWGRIGSENNTSCHRPVCRKEGVLLDYSKGGITVLLHEMDPFQKYS
VRHDYILLPEGALTNTRLRWQPFVINSGLVYSGVERAQWALDNLILGGAENIPSQL
VDTFDEGSHENFSFYNAVRTAGFCNPSFHLVWPNKKDKTNALSSRELITQP
GYMOPFKIVVGEATSCGDLHSMLEYTKDARSDWOLVOTOCPLPSSNSICGSPQF
HEATLYNAVNSWKRITIQLPDHVSSATQPRWIKGEBTEKQSWAIDHVIYGEACP
RLCSGHGYCTTAVCIDCSFQDDCSVFSEHLPSTIKONFESARVTEANWETIQGGA
IGSGGQLAPYAHGSLFXNGCOIQAATKPLDLTRASKIMFVLIQISTAAQDSCNSD
LSGPHTVDKAVILQYSVNNGITWYHIAQHPKDFTOAQRVSYNVLPLEARMKGYLLRW
QPRHNGTGHQWALDHVEVLVSTRQNTMMNFSRQHGLRHFYNNRRRSLRRYP"

BASE COUNT 2645 a 3029 c 2930 g 2583 t

Query Match 9.1%; Score 32; DB 10; Length 11187;
Best Local Similarity 100.0%; Pred. No. 6.1e-07;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 285 CGTGCAGTTCACGTGGAACAGACGACCTGC 316
|||||
Db 1315 CGTGCAGTTCACGTGGAACAGACGACCTGC 1346
|||||

RESULT 9
AX410790

LOCUS AX410790 11580 bp DNA linear PAT 14-JUN-2002

DEFINITION Sequence 3437 from Patent WO0229103.

ACCESSION AX410790

VERSION AX410790.1 GI:21443495

KEYWORDS human.

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.

TITLE Gene expression profiles in liver cancer

JOURNAL Patent: WO 0229103-A 3437 11-APR-2002;
GENE LOGIC INC (US)
FEATURES Location/Qualifiers
source 1..11580
/organism="Homo sapiens"
/db_xref="taxon:9606"
/notes="EMBL/GenBank Accession No. U79716"
BASE COUNT 3014 a 2696 c 2753 g 3116 t 1 others
ORIGIN

Query Match 7.4%; Score 26; DB 6; Length 11580;
Best Local Similarity 100.0%; Pred. No. 0.0019;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 208 GAGAAATTAGCCCTTCCAATGT 233
|||||
Db 1067 GAGAAATTAGCCCTTCCAATGT 1092
|||||

RESULT 10
HSU79716 11580 bp mRNA linear PRI 25-FEB-1997

LOCUS Human reelin (RELN) mRNA, complete cds.
DEFINITION U79716
ACCESSION U79716
VERSION U79716.1 GI:1743884

KEYWORDS Homo sapiens.

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 11580)

AUTHORS DeSilva,U., D'Arcangelo,G., Braden,V.V., Chen,J., Miao,G.G.,
Curran,T. and Green,E.D.

TITLE The human reelin gene: isolation, sequencing, and mapping on
chromosome 7

JOURNAL Genome Res. 7 (2), 157-164 (1997)

MEDLINE 97202106

PUBMED 9049633

REFERENCE 2 (bases 1 to 11580)

AUTHORS DeSilva,U., D'Arcangelo,G., Braden,V.V., Chen,J., Miao,G.G.,
Curran,T. and Green,E.D.

TITLE Direct Submission

JOURNAL Submitted (26-NOV-1996) National Center for Human Genome Research,
National Institutes of Health, 49 Convent Drive, MSC4431, Bethesda,
MD 20892, USA

FEATURES Location/Qualifiers
source 1..11580
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
gene 1..11580
/gene="RELN"
CDS 176..10558
/gene="RELN"
/codon_start=1
/product="reelin"
/protein_id="AAC51105.1"
/db_xref="GI:1743885"

translation="MERSGARQTFLLALLGATLRRARAAGYYPFRSPFFFLCTHHG
ELEGGEGEVLISLHIAQNPVYVPGQVHVITSTFTFDGLLVLTGLVTSVQASQ
SIGGSAFGFGIMSDHQFNQPMCSVASHVSHLPTNLSFIWIAPPAGTCGVNFAT
ATHRGQVIFDRLAQQLCEGAPTVTVHPHLAEIHSDSIILLRDDFDSTHQQLNPNI
WYFCNNECTGEQCGAIMHGNVATFCFYPRELITGLNTTASVLFQISGSGCRFS
YSDPSIIVLYAKNSADWLEKIRAPSNVSTIIHLYPEDAKAGENVQFOWKQENLR
VGEVYEAQWALDNILINSARHQRVLESLDPDVTGNLWLPFGATVVKHSCQDGNRIY
FHNSESENFATRDVLDSTEDIQOWSEEREESOPTGHDVLGAVTGTCSTIESGLS
MVFLKDGKRLCTPSSMTDITGNLRFYFMGGICDPGNSHENDIILYAFKLEGRKHIT
LDTLSSYSKVPFSLVSVINPELQTPATKFCRQKNHGHNRNVMAVDFHVLVPLPS
TMSHMIQFINLGCCTHQPNVSYLEFTNHNRSWLLHTECLPETCAGCPHLPSTVY
SSENTSGNRRITPLPNAALTNRNTRWQTGPIILGNMWAIDNVYIGPCLKFCSGRG
QCTRHGCKDCPDGSPACEMASOTPFMFISESGSRLSSVHNFTSIRCAEYFCGCV
LASGKALVFNKGRBLITSLFDSSQSRRFLOFTLRGSKSVLSTCRAPDQGEVLLH
YSYDNGTIVKLLLEHYSYLNTHPRIISVELPDDAKQFGIOFRWQPIHSTSGQEDVW

DEIIMTSVLNFSISLDFTNLVEVTSGLFYLGNVQPCYCGHDWTLCTGTGDSKLASSMRY
VETQSGIYASNYIOTSLVMGCGQKTYPHMDNQVLEYSTNHGLYTHLWQEBCLSPMP
SOEFTSASIVHASTOTWRVILVLPKQWSSATYFRWSSQYTYTAOWDALDSYIG
OCPNMCSCGCDGICRCDGOGYOGTECHPEAALPSTLMSDFENQNGWESDQWQVIG
GEIVPEQCGGYSGSSSLYFSKAKRQLVSWDLSTWDFVQFYIQIGESASCNKP
DSREGLVQYNNNGIOHLLAEMFYDSFSPREVIYELPAAATCTCFRFRWQPVF
SEYEDQWAAVDIIILSEKOKIIPINPTLPFNYKFAFDYPMQMSVWMLANEG
MYKNETFCAASIMIFGSGDGRFAVTRDLTKPGYVLQFKLNIQCAQFSTAPVL
LOYSHDAGSWFLVKEGCVYASAGCEGNSRELSEPTMYHTGDFEEMTRITVIPS
LASSKTRRWIOESSKQVNPFGDGVYISPCPSYCSGHGDCISGVCYCDLGYTAA
OCTCVSNVHNHMDREFEGLKSLPWYKITGAQVGTGCTLNDGKSLFNGPGKREAR
TYPLDTRNLVQFYIQIGKTSGLTICAPRTRNEGLVVOISDNGIMLHLLREDFM
SLEPOITIDLPQAKTAPATFRMWPQOGRKSAQALDDVLLIGNMSSQTFQDKF
DSGIDLQANNYIQGGVDIDCLMDTALIFTENICKPRIAETDFHVSASTFLQFEM
SMGCKPSNHSVQLOYSLNKGKDWHLVTECVPTIGLHYTSSIVTSERFNWK
RTVTLPLSTISPTFRFRIQANTYGVADSWAIDNVLASGCPWMSGKIGICDAGRCV
CDRGEGPCVCPVPLPSLLKDDFNGLHPDLWPVYGVAGERNLAGETIKSSTSLFK
GGLRLMSRDLDCNTMIVQSLRPIAKSTPERSHSLIQFSISGGITWHLMDDEFY
POTNITLFINVPLPOTATNTRFLWQYNNKKEEIMVDDFTIDGNVNNPVMLL
DTDFGPREDNFFYPGGNIGLYCPSKGAPEEDSAMVFNVEGHSITTRDLNVL
ENTIIQFEINVGSDSDSADPVREDFSRDFGATWHLPLCYHSSSHVSLCSTEHH
PSTVYAGTMQWRREVHFGKHLGCSVRERWOGFYFAGSQPVTAIDNVYIGQPC
EMKNGOGSCINGTKICDPYSGPTCKISTNPDLKDDFEGOLESDRELLMSGKP
SKCGLLSGNNLFFNEDGLRMLTDLDSHARFVQFMRGLGCGVDPKRSQPVLL
QYSLNGLSLQELFLSNSNVGRIYALEIPLKARSSTRLRWQSPSENGHYFSPW
VIDOILIGNISGNTVLEDDFTLDSRWLLHPGTTKMPVCGSTGDLVFIKASTRY
VYSTDAVNEDFSQIDFAASCSTDSVYAELEYSDLGLSWHPLVRDCLPTNVECS
RYHLORILVDTFNWTRITLPLPYTRSOATFRWHPAPFDKQOTWALDINVYIGDG
CIDMGSGHRCIOGNCVCDDEOGLYCDPETSPLTQLKDNFRAPSSONWITVNGGK
LSTVCGAVAGMALHFSGCCSRLVTVLDNLNAEYIQFYFMGILITPNNRNQGVLL
EYSVNGGITWLLMEIFYDOYSKPGFVNILLPPDAKEIATFRFRWQPRHGDQNDWA
IDNVLISGADQRTVMDTFSAPVQHERSPADAGPVGRIAFDMEDKTSVNEHWL
FHDDCTVERGSDPGMCLGSHDREYAVTHDLTPTEGIMWQFKRISYCKVSEKIA
QNOIHQVSTDFGVSNNVLPQCLPADPKCSVSQPSVEFFPTKWKRYITPLPESLV
GNPVRFRYQKYSDMQWADNVLGPGCLDNCRGHDCLEOCICDPGSGPNCYLTH
TLKTLKERDESEEIKPOLMNSLEGGSTCEGILAEADTALFYGGSTVRYQATQDLDL
RGAFLQYWRIGSENMTSCHRPICRKEGVLLDSTDGGITWTLHENDYOKYISVR
HDYILPEDALNTNRLRWQPFVINGVSVGERAQWALDNLIGGAEINPSQLVD
TFDEBTSHEENSFYPNAVRTAGFCGNPSFHYWPKKKDKTHNALSRELIIQPGY
MQQKIVGCEATSCGDLHSMLEYTKDARSWSQVLOTQCLPSSNSIGCSPPQFHE
ATLYNSVSSSKWRITIOLPDHVSSATOFRTWQKEETEKOSWALDHYIGEACPKL
CSGHGCTTGATCIDESEFGQDCCSVESHDLPYIKDNFESARVTEANWETIOGGVIG
SCGQLAPYAHGDSLYFNGCQIRQATKPLDLTRASKINFLVQIGSMSTDCNSDLS
GPHAVDKVLLQYVNNNGITHVIAHQHPQDFTQAOYRVSNNVPLEARMKGVLLRWQPL
RINGTGHQWALDHVEVLVSTRKQNNMNFGRQRLHFRYNNRRSLRRYP"

BASE COUNT 3014 a 2696 c 2753 g 3116 t 1 others

Query Match 7.4%; Score 26; DB 9; Length 11580;
Best Local Similarity 100.0%; Pred. No. 0.0019;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 208 GAGAAATAGAGCCCTTCCAATGT 233
|||||

Db 1067 GAGAAATAGAGCCCTTCCAATGT 1092

RESULT 11

G30936 11580 bp DNA linear STS 28-SEP-1998

LOCUS SW52926 Eric D. Green Homo sapiens STS genomic, sequence tagged

DEFINITION site.

ACCESSION G30936

VERSION G30936.1 GI:1923709

KEYWORDS STS.

SOURCE Homo sapiens.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 11580)

AUTHORS Bouffard,G.G., Iyer,L.M., Idol,J.R., Braden,V.V., Cunningham,A.F., Weintraub,L.A., Mohr-Tidwell,R.M., Peluso,D.C., Fulton,R.S., Leckie,M.P. and Green,E.D.

TITLE A collection of 1814 human chromosome 7-specific STSs
JOURNAL Genome Res. 7 (1), 59-64 (1997)
MEDLINE 97189344
PUBMED 9037602
REFERENCE 2 (bases 1 to 11580)
AUTHORS Green,E.D.
TITLE Human chromosome 7 STSs (1997)
JOURNAL Unpublished (1997)
COMMENT On Apr 3, 1997 this sequence version replaced gi:1706935.
SYNOPSIS: RELN
GDB_DSEC: RELN
Contact: Eric D. Green
Genome Technology Branch
National Human Genome Research Institute/NIH
49 Convent Dr., MSC4431, Bldg. 49, Rm. 2A08, Bethesda, MD 20892
Tel: 3014020201
Fax: 3014024735
Email: egreen@nhgri.nih.gov
Primer A: CTCCTACCTCTCTGAGGACGCCAA
Primer B: AAGCAGTTGCTGTGTCCACTG
STS size: 183
PCR Profile:
Prisoak: 0 degrees C for 0.00 minute(s)
Denaturation: 92 degrees C for 0.17 minute(s)
Annealing: 55 degrees C for 1.00 minute(s)
Polymerization: 72 degrees C for 1.00 minute(s)
PCR Cycles: 35
Thermal Cycler: PerkinElmer 9600
Protocol:
Template: 30-100 ng
Primer: each 1 uM
dNTPs: each 200 uM
Tag Polymerase: 0.05 units/uL
Total Vol: 10 uL
Buffer:
MgCl2: 1.5 mM
KCl: 100 mM
Tris-HCl: 10 mM
NH4Cl: 5 mM
pH: 8.6

This STS was developed from sequence determined by another investigator. See GenBank record: U79716 For additional information about the NHGRI chromosome 7 mapping project, see <http://www.nhgri.nih.gov/DIR/STB/CHR7>. Also see Genomics 11:548-64 (1991) [MUID=92128937].

FEATURES
source
Location/Qualifiers
1..11580
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="7"
/clone_lib="Eric D. Green"
1..11580
/gene="RELN"
1112..1294
/gene="RELN"
1112..1134
/gene="RELN"
primer_bind
primer_bind complement(1272..1294)
BASE COUNT 3014 a 2696 c 2753 g 3116 t 1 others
ORIGIN

Query Match 7.4%; Score 26; DB 11; Length 11580;
Best Local Similarity 100.0%; Pred. No. 0.0019;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 208 GAGAAATAGAGCCCTTCCAATGT 233
|||||

Db 1067 GAGAAATAGAGCCCTTCCAATGT 1092

RESULT 12

* overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1
 692 791: gap of 100 bp
 792 1481: contig of 690 bp in length
 1482 1581: gap of 100 bp
 1582 2255: contig of 714 bp in length
 2296 2395: gap of 100 bp
 2396 3014: contig of 619 bp in length
 3015 3114: gap of 100 bp
 3115 3850: contig of 736 bp in length
 3851 3950: gap of 100 bp
 3951 4619: contig of 669 bp in length
 4620 4719: gap of 100 bp
 4720 5407: contig of 688 bp in length
 5408 5507: gap of 100 bp
 5508 6196: contig of 689 bp in length
 6197 6296: gap of 100 bp
 6297 7009: contig of 713 bp in length
 7010 7109: gap of 100 bp
 7110 7805: contig of 696 bp in length
 7806 7905: gap of 100 bp
 7906 8615: contig of 710 bp in length
 8616 8715: gap of 100 bp
 8716 9429: contig of 714 bp in length
 9430 9529: gap of 100 bp
 9530 10249: contig of 720 bp in length
 10250 10349: gap of 100 bp
 10350 11059: contig of 710 bp in length
 11060 11159: gap of 100 bp
 11160 11856: contig of 697 bp in length
 11857 11956: gap of 100 bp
 11957 12632: contig of 676 bp in length
 12633 12732: gap of 100 bp
 12733 13425: contig of 693 bp in length
 13426 13525: gap of 100 bp
 13526 14193: contig of 668 bp in length
 14194 14293: gap of 100 bp
 14294 14959: contig of 666 bp in length
 14960 15059: gap of 100 bp
 15060 15752: contig of 693 bp in length
 15753 15852: gap of 100 bp
 15853 16545: contig of 693 bp in length
 16546 16645: gap of 100 bp
 16646 17336: contig of 691 bp in length
 17337 17436: gap of 100 bp
 17437 18129: contig of 693 bp in length
 18130 18229: gap of 100 bp
 18230 18944: contig of 715 bp in length
 18945 19044: gap of 100 bp
 19045 19763: contig of 719 bp in length
 19764 19863: gap of 100 bp
 19864 20586: contig of 723 bp in length
 20587 20686: gap of 100 bp
 20687 21421: contig of 735 bp in length
 21422 21521: gap of 100 bp
 21522 22211: contig of 690 bp in length
 22212 22311: gap of 100 bp
 22312 22962: contig of 651 bp in length
 22963 23062: gap of 100 bp
 23063 23748: contig of 686 bp in length
 23749 23848: gap of 100 bp
 23849 24535: contig of 687 bp in length
 24536 24635: gap of 100 bp
 24636 25329: contig of 694 bp in length
 25330 25429: gap of 100 bp
 25430 26129: contig of 700 bp in length
 26130 26229: gap of 100 bp
 26230 26929: contig of 700 bp in length
 26930 27029: gap of 100 bp

27030 27732: contig of 703 bp in length
 27733 27832: gap of 100 bp
 27833 28548: contig of 716 bp in length
 28549 28648: gap of 100 bp
 28649 29367: contig of 719 bp in length
 29368 29467: gap of 100 bp
 29468 30168: contig of 701 bp in length
 30169 30268: gap of 100 bp
 30269 30961: contig of 693 bp in length
 30962 31061: gap of 100 bp
 31062 31759: contig of 698 bp in length
 31760 31859: gap of 100 bp
 31860 32555: contig of 696 bp in length
 32556 32655: gap of 100 bp
 32656 33321: contig of 666 bp in length
 33322 33421: gap of 100 bp
 33422 34116: contig of 695 bp in length
 34117 34216: gap of 100 bp
 34217 34907: contig of 691 bp in length
 34908 35007: gap of 100 bp
 35008 35706: contig of 699 bp in length
 35707 35806: gap of 100 bp
 35807 36516: contig of 710 bp in length
 36517 36616: gap of 100 bp
 36617 37312: contig of 696 bp in length
 37313 37412: gap of 100 bp
 37413 38141: contig of 729 bp in length
 38142 38241: gap of 100 bp
 38242 38967: contig of 726 bp in length
 38968 39067: gap of 100 bp
 39068 39760: contig of 693 bp in length
 39761 39860: gap of 100 bp
 39861 40551: contig of 691 bp in length
 40552 40651: gap of 100 bp
 40652 41332: contig of 681 bp in length
 41333 41432: gap of 100 bp
 41433 42115: contig of 683 bp in length
 42116 42215: gap of 100 bp
 42216 42906: contig of 691 bp in length
 42907 43006: gap of 100 bp
 43007 43718: contig of 712 bp in length
 43719 43818: gap of 100 bp
 43819 44518: contig of 700 bp in length
 44519 44618: gap of 100 bp
 44619 45321: contig of 703 bp in length
 45322 45421: gap of 100 bp
 45422 46126: contig of 705 bp in length
 46127 46226: gap of 100 bp
 46227 46946: contig of 720 bp in length
 46947 47046: gap of 100 bp
 47047 47738: contig of 692 bp in length
 47739 47838: gap of 100 bp
 47839 48523: contig of 685 bp in length
 48524 48623: gap of 100 bp
 48624 49319: contig of 696 bp in length
 49320 49419: gap of 100 bp
 49420 50106: contig of 687 bp in length
 50107 50206: gap of 100 bp
 50207 50888: contig of 682 bp in length
 50889 50988: gap of 100 bp
 50989 51681: contig of 693 bp in length
 51682 51781: gap of 100 bp
 51782 52476: contig of 695 bp in length
 52477 52576: gap of 100 bp
 52577 53300: contig of 724 bp in length
 53301 53400: gap of 100 bp
 53401 54093: contig of 693 bp in length
 54094 54193: gap of 100 bp
 54194 54893: contig of 700 bp in length
 54894 54993: gap of 100 bp

Query Match
 Best Local Similarity

6.8%; Score 24; DB 2; Length 70905;
 100.0%; Pred. No. 0.027;

Matches	24;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	99	ATCTGTCTCTCCAGTTTCCCA	TGG	122					
Db	60522	ATCTGTCTCTCCAGTTTCCCA	TGG	60499					
RESULT 14									
AL590143									
LOCUS									
DEFINITION		Human DNA sequence from clone RP11-132M7 on chromosome 6, complete sequence.							
ACCESSION		AL590143							
VERSION		AL590143.4							
KEYWORDS		HTG.							
SOURCE		human.							
ORGANISM		Homo sapiens							
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
AUTHORS		1 (bases 1 to 73070)							
TITLE		Direct Submission							
JOURNAL		Submitted (10-MAY-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk							
COMMENT		requests: clonerequest@sanger.ac.uk							
		On May 14, 2001 this sequence version replaced gi:13751595.							
		During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.							
		This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at							
		http://www.sanger.ac.uk/Projects/C-elegans/wormpep/							
		This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at							
		http://www.sanger.ac.uk/HGP/Chr6							
		RP11-132M7 is from the library RPC1-11.1 constructed by the group of Pieter de Jong. For further details see							
		http://www.chori.org/bacpac/home.htm							
		VECTOR: pBACe3.6							
		IMPORTANT: This sequence is not the entire insert of clone RP11-132M7. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.							
		The true left end of clone RP1-3L1 is at 72971 in this sequence.							
		The true right end of clone RP11-533L1 is at 100 in this sequence.							
FEATURES		Location/Qualifiers							
source		1..73070							
		/organism="Homo sapiens"							
		/db_xref="taxon:9606"							
		/chromosome="6"							
		/clone="RP11-132M7"							
		/clone.lib="RPC1-11.1"							
repeat_region		41..418							
		/note="MLT1A1 repeat: matches 8..365 of consensus"							
repeat_region		853..1165							
		/note="MLT1A1 repeat: matches 1..321 of consensus"							
repeat_region		4023..4373							
		/note="MLT1A1 repeat: matches 1..362 of consensus"							
repeat_region		5311..5463							
		/note="MER58A repeat: matches 72..224 of consensus"							
repeat_region		5533..5743							

repeat_region /note="Tigger3(Golem) repeat: matches 1. .80 of consensus"
30330. .30652
repeat_region /note="LIME2 repeat: matches 5817. .6287 of consensus"
30812. .31142
repeat_region /note="AlusX repeat: matches 1. .303 of consensus"
32108. .32297
repeat_region /note="L1 repeat: matches 2822. .3028 of consensus"
32300. .33017
repeat_region /note="MER50 repeat: matches 1. .711 of consensus"
33115. .33682
repeat_region /note="L1 repeat: matches 3078. .3649 of consensus"
33789. .34325
repeat_region /note="MLT2B repeat: matches 2. .448 of consensus"
34391. .34544
repeat_region /note="L1 repeat: matches 3828. .3985 of consensus"
34545. .34856
repeat_region /note="AluY repeat: matches 1. .311 of consensus"
34857. .35048
repeat_region /note="L1 repeat: matches 3985. .4164 of consensus"
35415. .35601
repeat_region /note="MIR repeat: matches 21. .200 of consensus"
35701. .35836
repeat_region /note="L2 repeat: matches 2585. .2735 of consensus"
37327. .38045
repeat_region /note="L2 repeat: matches 2585. .2745 of consensus"
39417. .39539
repeat_region /note="MIR repeat: matches 19. .147 of consensus"
40146. .40185
repeat_region /note="20 copies 2 mer gt 85% conserved"
41086. .41217
repeat_region /note="L2 repeat: matches 1421. .1557 of consensus"
41236. .41369
repeat_region /note="MER63 repeat: matches 1. .766 of consensus"
41831. .42070
repeat_region /note="L1M4C repeat: matches 1617. .1860 of consensus"
42121. .42155
repeat_region /note="L1P4 repeat: matches 6112. .6146 of consensus"
42156. .42656
repeat_region /note="L1M4C repeat: matches 1072. .1565 of consensus"
43386. .43486
repeat_region /note="L1M49 repeat: matches 6210. .6308 of consensus"
45345. .45536
repeat_region /note="MIR repeat: matches 55. .246 of consensus"
46312. .46359
repeat_region /note="24 copies 2 mer ta 81% conserved"
48178. .48223
repeat_region /note="23 copies 2 mer ac 100% conserved".
48644. .48746
repeat_region /note="L2 repeat: matches 2159. .2262 of consensus"
49582. .49887
repeat_region /note="Alusq repeat: matches 1. .299 of consensus"
49888. .56031
repeat_region /note="L1P4 repeat: matches 4. .6144 of consensus"
56756. .56848
repeat_region /note="L2 repeat: matches 2598. .2700 of consensus"
58357. .58412
repeat_region /note="28 copies 2 mer tg 87% conserved"
58586. .58669
repeat_region /note="MIR repeat: matches 4. .92 of consensus"
58750. .58999
repeat_region /note="MIR repeat: matches 6. .252 of consensus"
60707. .60846
repeat_region /note="MER3 repeat: matches 53. .169 of consensus"

Query Match : 6.0%; Score 21; DB 9; Length 73070;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCAGTGTGGCACCATCATG 21
|||||
Db 2245 GAGCAGTGTGGCACCATCATG 2265

RESULT 15
AC015229
LOCUS Drosophila melanogaster, 98873 bp DNA linear HTG 16-NOV-1999
DEFINITION pieces. *** SEQUENCING IN PROGRESS ***, in ordered
AC015229
AC015229.1 GI:6436106
VERSION HTG; HTGS_PHASE2.
KEYWORDS Drosophila melanogaster.
SOURCE Drosophila melanogaster.
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 98873)
AUTHORS Adams, M. and Venter, J.C.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
COMMENT This sequence was identified as CDM:10214184 by the submitter.
For further information on this sequence e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
source
1. .98873
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
BASE COUNT 29559 a 19891 c 19792 g 29631 t
ORIGIN
Query Match 6.0%; Score 21; DB 2; Length 98873;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 188 CCGCTGATTGGATTCAGCTGG 208
|||||
Db 57715 CCGCTGATTGGATTCAGCTGG 57735
Search completed: November 6, 2002, 17:40:29
Job time : 3116 secs

SQ Sequence 7215 BP; 2335 A; 1369 C; 1559 G; 1952 T; 0 other;
 Query Match 5.4%; Score 19; DB 22; Length 7215;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 286 GTGCAGTTCGACGTGGAAC 304
 |||||
 DB 3906 GTGCAGTTCGACGTGGAAC 3924

RESULT 14
 ABL40174
 ID ABL40174 standard; DNA; 18 BP.
 XX
 AC ABL40174;
 XX
 DT 21-MAY-2002 (first entry)
 XX

Mouse reelin protein CR-50 epitope region PCR primer SEQ ID NO:11.

Mouse; reelin protein CR-50 epitope region; elucidation: neuron;
 cerebral disturbance; reelin protein; neuroprotective; PCR primer: ss.
 OS Mus musculus.
 XX
 PN JP2002017361-A.
 XX
 PD 22-JAN-2002.
 XX
 PF 04-JUL-2000; 2000JP-0202801.
 XX
 PR 04-JUL-2000; 2000JP-0202801.
 XX
 PA (RIKA) RIKAGAKU KENKYUSHO.
 XX
 DR WPI; 2002-221707/28.
 XX
 PT Reelin protein CR-50 epitope region, useful for diagnosis and treatment
 of cerebral disturbance -
 XX
 PS Example 2; Page 7; 16pp; Japanese.

CC The present invention describes the mouse reelin protein CR-50 epitope
 CC region, which contains the CR-50 antibody recognition site and is free
 CC from F-spondin domains and repetitive sites. Also described are: (1) an
 CC expression vector comprising a polynucleotide encoding a reelin protein
 CC epitope region; (2) host cells with transfected the expression vector;
 CC (3) polypeptides prepared by culture of the host cells; and (4)
 CC polynucleotides comprising the 351 base sequence given in ABL40165 which
 CC encodes the 117 amino acid sequence given in AB06244; and (5) use of
 CC the polynucleotide for diagnosis and/or treatment of diseases caused by
 CC abnormal positioning of neural cells, and stimulation of association of
 CC reelin protein. The mouse reelin protein CR-50 epitope region has
 CC neuroprotective activity, and can be used in the diagnosis and treatment
 CC of cerebral disturbance due to an abnormal reelin gene and positioning
 CC of neurons. The present sequence represents a PCR primer for the mouse
 CC reelin protein CR-50 epitope region, which is used in an example from
 CC the present invention.
 CC
 XX

SQ Sequence 18 BP; 4 A; 5 C; 6 G; 3 T; 0 other;

Query Match 5.1%; Score 18; DB 24; Length 18;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAGCAGTGTGACACATC 18
 |||||
 DB 1 GAGCAGTGTGACACATC 18

RESULT 15
 AAA70701

ID AAA70701 standard; cDNA; 975 BP.

XX AAA70701;

AC 11-DEC-2000 (first entry)

DE Human interleukin 6 receptor cDNA.

KM Human; interleukin-6 receptor; fungus; Pichia pastoris; PCR primer: ss;
 KM expression vector; immunoglobulin-like region; cytokine receptor region.
 XX
 OS Homo sapiens.

Key Location/Qualifiers
 FH 1..975
 FT CDS
 FT /tag= a
 FT /partial
 FT /product= "human IL-6R amino acids 20-344"
 FT /note= "no start or stop codon is given at the 5' or
 FT 3' ends of the sequence"
 FT /transl_except= (pos:169..171,aa:Glu)
 FT /transl_except= (pos:841..843,aa:Gly)

PN JP2000157280-A.

PD 13-JUN-2000.

PF 26-NOV-1998; 98JP-0335464.

PR 26-NOV-1998; 98JP-0335464.

PA (TOYD) TOSOH CORP.

DR WPI; 2000-468203/41.

DR P-PSDB; AAB15389.

PT Yeasts transformed with IL-6 receptor gene -

PS Example 1; Page 6-8; 10pp; Japanese.

CC The invention relates to the production of human interleukin-6 receptor
 CC (IL-6R) protein in the fungus Pichia pastoris. The fungus was
 CC transformed with an expression vector (pPIC9-A20HL) containing a gene
 CC encoding an IL-6R protein having an immunoglobulin-like region and
 CC cytokine receptor region and spanning amino acids from Leu20-Ala323. The
 CC human IL-6R region was PCR amplified using the primers AAA70702-A70703.
 CC This sequence represents the coding region for the IL-6R protein of the
 CC invention.
 CC
 XX

SQ Sequence 975 BP; 208 A; 289 C; 290 G; 188 T; 0 other;

Query Match 5.1%; Score 18; DB 21; Length 975;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 155 CCAGCATCACTGTGTCAAT 172
 |||||
 DB 762 CCAGCATCACTGTGTCAAT 779

Search completed: November 6, 2002, 16:48:34
 Job time : 314 secs

```
XX Key Location/Qualifiers
FH CDS 1..4593
FT /*tag= a
FT /partial
FT /product "Human protein, SEQ ID 6"
FT /note= "No start codon given"
XX
XX WO200127270-A1.
XX
XX 19-APR-2001.
XX
XX 25-AUG-2000; 2000MO-JP05756.
XX
XX 08-OCT-1999; 99JP-0288738.
XX 08-OCT-1999; 99JP-0288739.
XX 19-APR-2000; 2000JP-0123721.
XX
XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX
XX Taga T, Kimura N;
XX WPI; 2001-282027/29.
XX P-PSDB; AAB84882.
XX
XX DNA Sequences encoding YS68 protein involved in early hematopoiesis -
XX
XX Disclosure; Page 90-105; 245pp; Japanese.
XX
XX The present invention relates to novel DNA sequences encoding YS68
XX protein, which is involved in early hematopoiesis. The protein, YS68, is
XX involved in early haematopoiesis and the gene may be useful as a factor
XX for controlling it, or as a tool in development of related drugs. The
XX protein has potential use against tumours. The present sequence was used
XX in the present invention.
XX
XX Sequence 4883 BP; 1699 A; 938 C; 982 G; 1264 T; 0 other;
XX
XX Query Match 5.4%; Score 19; DB 22; Length 4883;
XX Best Local Similarity 100.0%; Pred. No. 14;
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 286 GTGCAGTTCAGTGGAAAC 304
XX |
XX Db 1574 GTGCAGTTCAGTGGAAAC 1592
XX
XX RESULT 12
XX AAH19497
XX ID AAH19497 standard; DNA: 7215 BP.
XX
XX AC AAH19497;
XX
XX 03-AUG-2001 (first entry)
XX
XX Human coding sequence, SEQ ID 13.
XX
XX Human; YS68 protein; early haematopoiesis; tumour; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 125..6925
XX FT /*tag= a
XX FT /product "Human protein, SEQ ID 14"
XX
XX WO200127270-A1.
XX
XX 19-APR-2001.
XX
XX 25-AUG-2000; 2000MO-JP05756.
XX
XX 08-OCT-1999; 99JP-0288738.
XX
```

```
PR 08-OCT-1999; 99JP-0288739.
PR 19-APR-2000; 2000JP-0123721.
XX
XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX
XX Taga T, Kimura N;
XX WPI; 2001-282027/29.
XX P-PSDB; AAB84885.
XX
XX DNA Sequences encoding YS68 protein involved in early hematopoiesis -
XX
XX Claim 1; Page 180-202; 245pp; Japanese.
XX
XX The present invention relates to novel DNA sequences encoding YS68
XX protein, which is involved in early haematopoiesis. The protein, YS68, is
XX involved in early haematopoiesis and the gene may be useful as a factor
XX for controlling it, or as a tool in development of related drugs. The
XX protein has potential use against tumours. The present sequence was used
XX in the present invention.
XX
XX Sequence 7215 BP; 2335 A; 1369 C; 1559 G; 1952 T; 0 other;
XX
XX Query Match 5.4%; Score 19; DB 22; Length 7215;
XX Best Local Similarity 100.0%; Pred. No. 14;
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 286 GTGCAGTTCAGTGGAAAC 304
XX |
XX Db 3906 GTGCAGTTCAGTGGAAAC 3924
XX
XX RESULT 13
XX AAH19498
XX ID AAH19498 standard; DNA: 7215 BP.
XX
XX AC AAH19498;
XX
XX 03-AUG-2001 (first entry)
XX
XX Human coding sequence, SEQ ID 15.
XX
XX Human; YS68 protein; early haematopoiesis; tumour; ss.
XX
XX Homo sapiens.
XX
XX WO200127270-A1.
XX
XX 19-APR-2001.
XX
XX 25-AUG-2000; 2000MO-JP05756.
XX
XX 08-OCT-1999; 99JP-0288738.
XX 08-OCT-1999; 99JP-0288739.
XX 19-APR-2000; 2000JP-0123721.
XX
XX (CHUG-) CHUGAI RES INST MOLECULAR MEDICINE INC.
XX
XX Taga T, Kimura N;
XX WPI; 2001-282027/29.
XX
XX DNA Sequences encoding YS68 protein involved in early hematopoiesis -
XX
XX Disclosure; Page 220-242; 245pp; Japanese.
XX
XX The present invention relates to novel DNA sequences encoding YS68
XX protein, which is involved in early haematopoiesis. The protein, YS68, is
XX involved in early haematopoiesis and the gene may be useful as a factor
XX for controlling it, or as a tool in development of related drugs. The
XX protein has potential use against tumours. The present sequence was used
XX in the present invention.
XX
```

CC of neurons. The present sequence represents a PCR primer for the mouse
CC reelin protein CR-50 epitope region, which is used in an example from
CC the present invention.

XX
SQ Sequence 22 BP; 5 A; 9 C; 6 G; 2 T; 0 other;

Query Match 6.3%; Score 22; DB 24; Length 22;

Best Local Similarity 100.0%; Pred. No. 0.47;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 330 GTATGAGCCTGCTGGCCCTG 351

DB 22 GTATGAGCCTGCTGGCCCTG 1

RESULT 9

ABL40173

ID ABL40173 standard; DNA: 21 BP.

XX
XX ABL40173;

XX 21-MAY-2002 (first entry)

DE Mouse reelin protein CR-50 epitope region PCR primer SEQ ID NO:10.

XX
XX Mouse: reelin protein CR-50 epitope region; elucidation; neuron;

KW cerebral disturbance; reelin protein; neuroprotective; PCR primer; ss.

XX
OS Mus musculus.

XX
PN JP2002017361-A.

XX
PD 22-JAN-2002.

XX
PF 04-JUL-2000: 2000JP-0202801.

XX
PR 04-JUL-2000: 2000JP-0202801.

XX
PA (RIKA) RIKAGAKU KENKYUSHO.

XX
DR WPI: 2002-221707/28.

XX
PT Reelin protein CR-50 epitope region, useful for diagnosis and treatment
of cerebral disturbance -

XX
PS Example 1; Page 6; 16pp; Japanese.

CC The present invention describes the mouse reelin protein CR-50 epitope
region, which contains the CR-50 antibody recognition site and is free
from F-spondin domains and repetitive sites. Also described are: (1) an
expression vector comprising a polynucleotide encoding a reelin protein
epitope region; (2) host cells with transfected the expression vector;
CC (3) polypeptides prepared by culture of the host cells; and (4)
CC polynucleotides comprising the 351 base sequence given in ABL40165 which
CC encodes the 117 amino acid sequence given in AB06244; and (5) use of
CC the polynucleotide for diagnosis and/or treatment of diseases caused by
CC abnormal positioning of neural cells, and stimulation of association of
CC reelin protein. The mouse reelin protein CR-50 epitope region has
CC neuroprotective activity, and can be used in the diagnosis and treatment
CC of cerebral disturbance due to an abnormal reelin gene and positioning
CC of neurons. The present sequence represents a PCR primer for the mouse
CC reelin protein CR-50 epitope region, which is used in an example from
CC the present invention.

XX
SQ Sequence 21 BP; 5 A; 5 C; 7 G; 4 T; 0 other;

Query Match 6.0%; Score 21; DB 24; Length 21;

Best Local Similarity 100.0%; Pred. No. 1.5;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCAGTGTGGCACCATCATG 21

DB 1 GAGCAGTGTGGCACCATCATG 21

RESULT 10
AAD22777
ID AAD22777 standard; DNA: 20 BP.

XX
AC AAD22777;

XX
DT 26-FEB-2002 (first entry)

DE Mouse truncated reelin DNA existence detecting RL-11 RACE primer.

XX
XX Mouse: reelin; F-spondin domain; CR-50 epitope; gene therapy; agyria;

KW RACE; rapid amplification of cDNA end; polymicrogyria; PCR primer;

XX
XX ectopic gray matter; ss.

XX
PN BP149844-A2.

XX
PD 31-OCT-2001.

XX
PF 11-APR-2001: 2001EP-0303411.

XX
PR 11-APR-2000: 2000JP-0109954.

XX
PA (RIKE) RIKEN KK.

XX
PI Mikoshiba K, Tabata H, Nakajima K;

XX
DR WPI: 2002-019320/03.

XX
PT Novel truncated Reelin protein containing F-spondin domain and CR-50
PT recognition site of Reelin protein, but not having Reelin repeat site,
PT useful to treat diseases including agyria due to abnormal neuron
PT alignment -

XX
PS Example 6; Page 8; 47pp; English.

XX
CC The invention relates to a truncated Reelin protein comprising a
CC F-spondin domain and a CR-50 recognition site but no reelin protein
CC repeat site. Reelin is an essential molecule in developing a normal
CC laminated structure of cerebrum. The truncated reelin protein and its
CC DNA are useful for treating diseases including agyria, polymicrogyria,
CC and ectopic gray matter due to abnormal neuronal alignment. Truncated
CC reelin protein DNA is useful in gene therapy. The present sequence is
CC a gene specific RACE (rapid amplification of cDNA end) PCR primer used
CC for detecting mouse truncated reelin DNA existence.

XX
SQ Sequence 20 BP; 4 A; 3 C; 7 G; 6 T; 0 other;

Query Match 5.7%; Score 20; DB 24; Length 20;

Best Local Similarity 100.0%; Pred. No. 4.8;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 191 CTGATTGGATTGACGTGGAG 210

DB 1 CTGATTGGATTGACGTGGAG 20

RESULT 11
AAH19492
ID AAH19492 standard; DNA: 4883 BP.

XX
AC AAH19492;

XX
DT 03-AUG-2001 (first entry)

XX
DE Human coding sequence, SEQ ID 5.

XX
KW Human: YS68 protein; early haematopoiesis; tumour; ss.

XX
OS Homo sapiens.

CC F-spondin at the N terminus, a stretch of positively charged amino
CC acids at the C terminus, and a series of eight internal repeats of
CC 350-390 amino acids, each repeat containing two related sub-domains
CC that flank a pattern of conserved cysteine residues known as an
CC epidermal growth factor (EGF)-like motif. (I) has neuroprotective,
CC neurotropic and antipapillary activities, and can be used as a modulator
CC of reelin-LDHR interaction. (I) is useful in screen for compounds that
CC modulate reelin binding to an LDHR, in an assay system, where the assay
CC system comprises a microplate array and an automated robotic
CC microprocessor controlled system for adding and removing reagents to
CC the microplate array. The compounds identified by the above screening
CC method are useful as therapeutic agents to provide or alleviate a
CC diverse spectrum of diseases including neurodegenerative disorders such
CC as Alzheimer's disease, to facilitate neuronal regeneration after
CC injury, to prevent or alleviate lipid metabolism diseases, to enhance
CC cognitive functions and memory or to ameliorate other developmental
CC disorders. The present sequence encodes human reelin, which is used in
CC the exemplification of the present invention.

SO Sequence 11580 BP; 3014 A; 2696 C; 2753 G; 3116 T; 1 other;

Query Match 7.4%; Score 26; DB 24; Length 11580;

Best Local Similarity 100.0%; Pred. No. 0.0041;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 208 GAGAAATTAGAGCCCTTCCATGT 233

DB 1067 GAGAAATTAGAGCCCTTCCATGT 1092

RESULT 7

AAS89484

ID AAS89484 standard; cDNA; 11632 BP.

XX AAS89484;

XX 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #25288.

KM Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PR 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR P-PSDB; ABG25297.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

PS Claim 1; SEQ ID No 25288; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.

SO Sequence 11632 BP; 3028 A; 2716 C; 2770 G; 3117 T; 1 other;

Query Match 7.4%; Score 26; DB 23; Length 11632;

Best Local Similarity 100.0%; Pred. No. 0.0041;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 208 GAGAAATTAGAGCCCTTCCATGT 233

DB 1067 GAGAAATTAGAGCCCTTCCATGT 1092

RESULT 8

ABL40169/C

ID ABL40169 standard; DNA; 22 BP.

XX ABL40169;

XX 21-MAY-2002 (first entry)

DE Mouse reelin protein CR-50 epitope region PCR primer SEQ ID NO:6.

KM Mouse; reelin protein CR-50 epitope region; elucidation; neuron;

KW cerebral disturbance; reelin protein; neuroprotective; PCR primer; ss.

OS Mus musculus.

PN JP2002017361-A.

PD 22-JAN-2002.

PE 04-JUL-2000; 2000JP-0202801.

PR 04-JUL-2000; 2000JP-0202801.

PA (RIKA) RIKAGAKU KENKYUSHO.

DR WPI; 2002-221707/28.

PI Reelin protein CR-50 epitope region, useful for diagnosis and treatment
PI of cerebral disturbance

PS Example 1; Page 6; 16pp; Japanese.

XX The present invention describes the mouse reelin protein CR-50 epitope
XX region, which contains the CR-50 antibody recognition site and is free
XX from F-spondin domains and repetitive sites. Also described are: (1) an
XX expression vector comprising a polynucleotide encoding a reelin protein
XX epitope region; (2) host cells with transfected the expression vector;
XX (3) polypeptides prepared by culture of the host cells; and (4)
XX polynucleotides comprising the 351 base sequence given in ABL40165 which
XX encodes the 117 amino acid sequence given in ABL40165 and (5) use of
XX the polynucleotide for diagnosis and/or treatment of diseases caused by
XX abnormal positioning of neural cells, and stimulation of association of
XX reelin protein. The mouse reelin protein CR-50 epitope region has
XX neuroprotective activity, and can be used in the diagnosis and treatment
XX of cerebral disturbance due to an abnormal reelin gene and positioning

RESULT 3
ABA92603 standard; cDNA: 11673 BP.
ID ABA92603
AC ABA92603;
XX
XX
DT 21-MAR-2002 (first entry)
XX
DE Mouse reelin encoding cDNA SEQ ID NO:3.
XX
XX Mouse; reelin: low density lipoprotein receptor; LDLR; neuroprotective;
KM extracellular glycoprotein; nontropic; antilipemic; Alzheimer's disease;
KM neurodegenerative disorder; neuronal regeneration; cognitive function;
KM lipid metabolism disease; memory; developmental disorder; gene; ss.
XX
XX Mus musculus.
XX
FH Key Location/Qualifiers.
FT CDS 283..10668
F /*tag= a
F /product= "mouse reelin"
XX
XX US6323177-B1.
XX
XX 27-NOV-2001.
XX
XX 16-JUN-1999; 99US-0334220.
XX
XX 16-JUN-1999; 99US-0334220.
XX
XX (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
XX Curran T, D'Arcangelo G;
XX
XX WPI: 2002-096596/13.
XX P-PSDB; ABB05008.
XX
XX Novel composition useful for screening compounds that modulate Reelin
PT binding to low density lipoprotein receptor, comprising an isolated
PT Reelin polypeptide and low density lipoprotein receptor -
XX
XX
XX Example 1: Column 65-74; 45pp: English.
XX
XX The present invention describes a composition (I) comprising an
CC isolated reelin protein (II) bound to an isolated low density lipoprotein
CC receptor (LDLR) (III). (II) is an extracellular glycoprotein of
CC approximately 385 kDa containing a small region of similarity with
CC F-spondin at the N terminus, a stretch of positively charged amino
CC acids at the C terminus, and a series of eight internal repeats of
CC 350-390 amino acids, each repeat containing two related sub-domains
CC that flank a pattern of conserved cysteine residues known as an
CC epidermal growth factor (EGF)-like motif. (I) has neuroprotective,
CC neurotropic and antilipemic activities, and can be used as a modulator
CC of reelin-LDLR interaction. (I) is useful in screen for compounds that
CC modulate reelin binding to an LDLR, in an assay system, where the assay
CC system comprises a microplate array and an automated robotic
CC microprocessor controlled system for adding and removing reagents to
CC the microplate array. The compounds identified by the above screening
CC method are useful as therapeutic agents to provide or alleviate a
CC diverse spectrum of diseases including neurodegenerative disorders such
CC as Alzheimer's disease, to facilitate neuronal regeneration after
CC injury, to prevent or alleviate lipid metabolism diseases, to enhance
CC cognitive functions and memory or to ameliorate other developmental
CC disorders. The present sequence encodes mouse (Mus musculus) reelin,
CC which is used in the exemplification of the present invention.
XX
XX Sequence 11673 BP; 2831 A; 2985 C; 2985 G; 2872 T; 0 other:
XX

Query Match 100.0%; Score 351; DB 24; Length 11673;
Best Local Similarity 100.0%; Pred. No. 1.4e-166;
Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGCAGTGTGGCACCACATCATGCAATGCTGTACCTTCTGTGAGCCGTACGGCCT 60
Dd |||||||
Dd 970 GAGCAGTGTGGCACCACATCATGCAATGCTGTACCTTCTGTGAGCCGTACGGCCT 1029
QY 61 CGAGAGCTGACACCATGCTGTGACACAAACACAGCATGCTGTGACATTTCCATT 120
Dd |||||||
Dd 1030 CGAGAGCTGACACCATGCTGTGACACAAACACAGCATGCTGTGACATTTCCATT 1089
QY 121 GGGTCAGATCATGTCGATTTAGTTACTGTGACCCGACATCACTGTGTATAGCCCAAG 180
Dd |||||||
Dd 1090 GGGTCAGATCATGTCGATTTAGTTACTGTGACCCGACATCACTGTGTATAGCCCAAG 1149
QY 181 AACCAATACCGTGTATGATTCACCTGAGAAAATTTAGAGCCCTTCCATGTAGACACA 240
Dd |||||||
Dd 1150 AACCAATACCGTGTATGATTCACCTGAGAAAATTTAGAGCCCTTCCATGTAGACACA 1209
QY 241 GTCATCCACATCCTGTACCTCCCGAGAGAACCCAAAGGGAGAGCGTGATTCAGCTG 300
Dd |||||||
Dd 1210 GTCATCCACATCCTGTACCTCCCGAGAGAACCCAAAGGGAGAGCGTGATTCAGCTG 1269
QY 301 AAACAGGACACGCTGCGAGTGGGTATGAGGCGCTGTGGGCCCTG 351
Dd |||||||
Dd 1270 AAACAGGACACGCTGCGAGTGGGTATGAGGCGCTGTGGGCCCTG 1320
XX
XX
XX RESULT 4
XX ABI9284
XX ID ABI9284 standard; cDNA: 11673 BP.
XX
XX
XX ABI9284;
XX
XX 07-MAR-2002 (first entry)
XX
XX Mouse ischaemic condition related cDNA sequence SEQ ID NO:128.
XX
XX
XX Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
KM vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.
XX
XX Mus musculus.
XX
XX PN WO200188188-A2
XX
XX 22-NOV-2001.
XX
XX 18-MAY-2001; 2001WO-JP04192.
XX
XX 18-MAY-2000; 2000JP-0145977.
XX
XX (UYNI-) UNIV NIHON SCHOOL JURIDICAL PERSON.
XX
XX Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
XX
XX WPI: 2002-034733/04.
XX P-PSDB; ABB57065.
XX
XX Examining the ischemic condition (e.g. occlusive ischemia) by measuring
PT expression levels of particular genes defined in the specification or
PT by determining the expression profile of a gene group comprising these
PT genes -
XX
XX Claim 2; Page 362-385; 2690pp: English.
XX
XX The present invention describes a method for examining ischaemic
CC conditions, comprising measuring the expression levels of particular
CC genes (I) in a test sample or determining the expression profile of a
CC gene group in the sample comprising genes selected from (I). The method
CC is useful for examining the ischaemic condition (e.g. compressive
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
CC expression levels of particular genes (ABI9284 to ABI9912), encoding
CC the protein sequences in ABB57020 to ABB57374) or by determining the
CC expression profile of a gene group comprising these genes. The
CC expression levels or expression profiles produced by these genes are
CC used as an indicator when screening for ischaemic condition-improving

DR WPI; 2002-221707/28.
 DR P-PSDB; ABB06244.
 XX Reelin protein CR-50 epitope region, useful for diagnosis and treatment
 PT of cerebral disturbance -
 PS Claim 5; Page 11; 16pp; Japanese.
 XX
 CC The present sequence encodes the mouse reelin protein CR-50 epitope
 CC region, which contains the CR-50 antibody recognition site and is free
 CC from F-spondin domains and repetitive sites. Also described are: (1) an
 CC expression vector comprising a polynucleotide encoding a reelin protein
 CC epitope region; (2) host cells with transfected the expression vector;
 CC (3) polypeptides prepared by culture of the host cells; and (4)
 CC polynucleotides comprising the 351 base sequence given in ABL40165 which
 CC encodes the 117 amino acid sequence given in ABB06244; and (5) use of
 CC the polynucleotide for diagnosis and/or treatment of diseases caused by
 CC abnormal positioning of neural cells, and stimulation of association of
 CC reelin protein. The mouse reelin protein CR-50 epitope region has
 CC neuroprotective activity, and can be used in the diagnosis and treatment
 CC of cerebral disturbance due to an abnormal reelin gene and positioning
 CC of neurons.
 XX
 SQ Sequence 351 BP; 86 A; 98 C; 92 G; 75 T; 0 other;
 Query Match 100.0%; Score 351; DB 24; Length 351;
 Best Local Similarity 100.0%; Pred. No. 1.5e-166;
 Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GACCACTGTGGCACCATCATGATGCGATCTGTCACCTTCTGTGAGCCGTACGGCCCT 60
 DB 1 GACCACTGTGGCACCATCATGATGCGATCTGTCACCTTCTGTGAGCCGTACGGCCCT 60
 QY 61 CGAGAGCTGACCCACCATGCTGAAACACAAACAGCATCTGCTCCAGTTTCCATT 120
 DB 61 CGAGAGCTGACCCACCATGCTGAAACACAAACAGCATCTGCTCCAGTTTCCATT 120
 QY 121 GGGTCAGAGATCATGCTGATTTAGTACTGTACCCAGCATCACTGTGTCAATACGCCAAG 180
 DB 121 GGGTCAGAGATCATGCTGATTTAGTACTGTACCCAGCATCACTGTGTCAATACGCCAAG 180
 QY 181 AACAAATACCCGCTGATTTAGTACTGTACCCAGCATCACTGTGTCAATACGCCAAG 240
 DB 181 AACAAATACCCGCTGATTTAGTACTGTACCCAGCATCACTGTGTCAATACGCCAAG 240
 QY 241 GTCATCCACATCTCTGTACCTCCCGAGAGCAAAAGGGAGAGCCGTGCAGTTCCAGTGG 300
 DB 241 GTCATCCACATCTCTGTACCTCCCGAGAGCAAAAGGGAGAGCCGTGCAGTTCCAGTGG 300
 QY 301 AACAGAGACAGCCTGGAGTGGGTATGAGGCTGTGAGCCCTG 351
 DB 301 AACAGAGACAGCCTGGAGTGGGTATGAGGCTGTGAGCCCTG 351
 RESULT 2
 AAD22754
 ID AAD22754 standard; cDNA; 2745 BP.
 XX
 AC AAD22754;
 XX
 DT 26-FEB-2002 (first entry)
 XX
 DE Mus musculus truncated reelin cDNA.
 XX
 KM Mouse; reelin; F-spondin domain; CR-50 epitope; gene therapy; agyria;
 XX polynucleotides; ectopic gray matter; ss.
 OS Mus musculus.
 XX
 FH Key Location/Qualifiers
 FT CDS 283..2052
 FT /tag= a
 FT /product= "Mouse truncated reelin protein"

FT sig_peptide 283..363
 FT /tag= b
 FT mat_peptide 364..2049
 FT /tag= c
 FT misc_feature /product= "Mature truncated reelin protein"
 FT /tag= d
 FT /note= "Encodes F-spondin domain"
 FT misc_feature 970..1320
 FT /tag= e
 FT /note= "Encodes CR-50 epitope region"
 XX
 /PN BP1149844-A2.
 XX
 PD 31-OCT-2001.
 XX
 PE 11-APR-2001; 2001BP-0303411.
 XX
 PR 11-APR-2000; 2000BP-0109954.
 XX
 RA (Riken) RIKEN KK.
 XX
 PI Mikoshiba K, Tabata H, Nakajima K;
 XX
 DR WPI; 2002-019320/03.
 XX
 DR P-PSDB; AAE13606.
 XX
 PT Novel truncated Reelin protein containing F-spondin domain and CR-50
 PT recognition site of Reelin protein, but not having Reelin repeat site,
 PT useful to treat diseases including agyria due to abnormal neuron
 PT alignment -
 PS
 XS Claim 10; Page 20-26; 47pp; English.
 CC The invention relates to a truncated Reelin protein comprising a
 CC F-spondin domain and a CR-50 recognition site but no reelin protein
 CC repeat site. Reelin is an essential molecule in developing a normal
 CC laminated structure of cerebrum. The truncated reelin protein and its
 CC DNA are useful for treating diseases including agyria, polymicrogyria,
 CC and ectopic gray matter due to abnormal neuronal alignment. Truncated
 CC reelin protein DNA is useful in gene therapy. The present sequence is
 CC a cDNA encoding Mus musculus truncated reelin protein.
 XX
 SQ Sequence 2745 BP; 661 A; 716 C; 714 G; 654 T; 0 other;
 Query Match 100.0%; Score 351; DB 24; Length 2745;
 Best Local Similarity 100.0%; Pred. No. 1.5e-166;
 Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GACCACTGTGGCACCATCATGATGCGATCTGTCACCTTCTGTGAGCCGTACGGCCCT 60
 DB 970 GACCACTGTGGCACCATCATGATGCGATCTGTCACCTTCTGTGAGCCGTACGGCCCT 1029
 QY 61 CGAGAGCTGACCCACCATGCTGAAACACAAACAGCATCTGCTCCAGTTTCCATT 120
 DB 1030 CGAGAGCTGACCCACCATGCTGAAACACAAACAGCATCTGCTCCAGTTTCCATT 1089
 QY 121 GGGTCAGAGATCATGCTGATTTAGTACTGTACCCAGCATCACTGTGTCAATACGCCAAG 180
 DB 1090 GGGTCAGAGATCATGCTGATTTAGTACTGTACCCAGCATCACTGTGTCAATACGCCAAG 1149
 QY 181 AACAAATACCCGCTGATTTAGTACTGTACCCAGCATCACTGTGTCAATACGCCAAG 240
 DB 1150 AACAAATACCCGCTGATTTAGTACTGTACCCAGCATCACTGTGTCAATACGCCAAG 1209
 QY 241 GTCATCCACATCTCTGTACCTCCCGAGAGCAAAAGGGAGAGCCGTGCAGTTCCAGTGG 300
 DB 1210 GTCATCCACATCTCTGTACCTCCCGAGAGCAAAAGGGAGAGCCGTGCAGTTCCAGTGG 1269
 QY 301 AACAGAGACAGCCTGGAGTGGGTATGAGGCTGTGAGCCCTG 351
 DB 1270 AACAGAGACAGCCTGGAGTGGGTATGAGGCTGTGAGCCCTG 1320

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 6, 2002, 15:27:06 ; Search time 300 Seconds
(without alignments)
2634.838 Million cell updates/sec

Title: US-09-897-438b-1

Sequence: 1 gagcagctggcaccatcat.....atgagcctcgcggcgcctg 351

Scoring table: OLIGO_NUC
Gapex 60.0 , Capext 60.0

Searched: 2185239 seqs, 112599159 residues

W size: 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- 1: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SID22/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SID22/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	351	100.0	351	24	ABL40165	Mouse reelin prote
2	351	100.0	2745	24	AAD22754	Mus musculus trunc
3	351	100.0	11673	24	ABA92603	Mouse reelin enco
4	351	100.0	11673	24	ABA92604	Mouse reelin enco
5	26	7.4	11580	24	ABN96939	Gene #3437 used to
6	26	7.4	11580	24	ABN96939	Human reelin enco
7	26	6.3	11632	23	ABA59484	DNA encoding novel
8	22	6.0	21	24	ABA40169	Mouse reelin prote
9	21	6.0	21	24	ABA40173	Mouse reelin prote

10	20	5.7	20	24	AAD22777	Mouse truncated re
11	19	5.4	4883	22	AAH19492	Human coding seque
12	19	5.4	7215	22	AAH19497	Human coding seque
13	19	5.4	7215	22	AAH19498	Human coding seque
14	18	5.1	18	24	ABL40174	Mouse reelin prote
15	18	5.1	975	21	AAV07071	Human interleukin
16	18	5.1	1035	21	AAV07071	Human interleukin
17	18	5.1	1074	18	AAV04440	SR345 coding seque
18	18	5.1	1260	20	AAV02922	Interleukin 6 rece
19	18	5.1	1466	19	AAV60296	Human IL-6 recepto
20	18	5.1	1545	21	AAV07073	Human interleukin-
21	18	5.1	1627	18	AAV07073	IL-6R/IL-6 fusion
22	18	5.1	2061	10	AAV03440	Human fusion poly
23	18	5.1	2066	14	AAV03440	Sequence encoding
24	18	5.1	2067	10	AAV03440	IL-6 receptor codi
25	18	5.1	2087	10	AAV0847	DNA contg. region
26	18	5.1	3319	17	AAV33369	Human secreted pro
27	18	5.1	3319	17	AAV33369	Interleukin-6 rece
28	18	5.1	3319	19	AAV60295	Human interleukin-
29	18	5.1	3319	21	AAV33364	Human low adenosin
30	18	5.1	3319	21	AAV33364	Human low adenosin
31	18	5.1	3477	21	AAV09047	Human adenosine re
32	18	5.1	3507	21	AAV09046	Human CDNA differe
33	18	5.1	4513	21	AAV21365	Fusion polypeptide
34	18	5.1	4873	21	AAV35243	Human low adenosin
35	18	5.1	9720	24	ABD56176	Human adenosine re
36	18	5.1	198285	24	ABR84659	Hordeum vulgare va
37	18	5.1	198285	24	ABR84659	Human CDNA differe
38	17	4.8	22	21	AAC63485	Gene #3817 used to
39	17	4.8	47	18	AAV90825	Beta-actin gene PC
40	17	4.8	325	22	AAV36445	Bacillus steatother
41	17	4.8	335	18	AAV13260	Human musculuskele
42	17	4.8	429	24	ABN95973	N. tabacum strain
43	17	4.8	559	22	ABA63804	Gene #2471 used to
44	17	4.8	559	22	ABA30992	Human foetal liver
45	17	4.8	559	22	AAK12314	Probe #9458 for ge
						Human brain expres

ALIGNMENTS

RESULT 1	ABL40165	standard; DNA: 351 BP.
ID	ABL40165	
AC	ABL40165;	
XX		
DT	21-MAY-2002	(first entry)
XX		
DE	Mouse reelin protein CR-50 epitope region encoding DNA SEQ ID NO:1.	
XX		
KW	Mouse; reelin protein CR-50 epitope region; elucidation; neuron;	
KM	cerebral disturbance; reelin protein; neuroprotective; gene; ds.	
XX		
OS	Mus musculus.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	1..351
FT		/*tag= a
FT		/partial
FT		/product= "reelin protein CR-50 epitope region"
FT		/note= "no start or stop codons given"
PN	JP2002017361-A.	
XX		
PD	22-JAN-2002.	
XX		
PF	04-JUL-2000; 2000JP-0202801.	
XX		
PR	04-JUL-2000; 2000JP-0202801.	
XX		
PA	(RIKA) RIKAGAKU KENKYUSHO.	
XX		

Query Match 4.8%; Score 17; DB 2; Length 47;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 267 GGAAGCCAAAGGGGAGA 283
Db 17 GGAAGCCAAAGGGGAGA 33

RESULT 13
US-09-480-921B-9/C

; Sequence 9, Application US/09480921B
; Patent No. 6387637
; GENERAL INFORMATION:
; APPLICANT: Levin, Joshua Z.
; APPLICANT: Budziszewski, Gregory J.
; APPLICANT: Potter, Sharon L.
; APPLICANT: Weglich, Lynette M.
; TITLE OF INVENTION: Herbicide Target Genes and Methods
; CURRENT FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1434
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1434)
US-09-480-921B-9

Query Match 4.8%; Score 17; DB 4; Length 1434;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 195 TTGGATTGAGCTGAGA 211
Db 38 TTGGATTGAGCTGAGA 22

RESULT 14
US-09-480-921B-28/C
; Sequence 28, Application US/09480921B
; Patent No. 6387637
; GENERAL INFORMATION:
; APPLICANT: Levin, Joshua Z.
; APPLICANT: Budziszewski, Gregory J.
; APPLICANT: Potter, Sharon L.
; APPLICANT: Weglich, Lynette M.
; TITLE OF INVENTION: Herbicide Target Genes and Methods
; FILE REFERENCE: PB/5-30780A
; CURRENT APPLICATION NUMBER: US/09/480,921B
; CURRENT FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 1434
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1434)
; OTHER INFORMATION: encodes SEQ ID NO:29
US-09-480-921B-28

Query Match 4.8%; Score 17; DB 4; Length 1434;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 195 TTGGATTGAGCTGAGA 211

Db 38 TTGGATTGAGCTGAGA 22

RESULT 15
US-09-108-010B-11
; Sequence 11, Application US/09108010B
; Patent No. 6362399
; GENERAL INFORMATION:
; APPLICANT: ANTHONY JOHN KINNEY
; GARY MICHAEL FADER
; TITLE OF INVENTION: SUPPRESSION OF SPECIFIC
; CLASSES OF SOYBEAN SEED
; PROTEIN GENES
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/108,010B
; FILING DATE: 30-Jun-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/019,940
; FILING DATE: JUNE 14, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: LYNNE M. CHRISTENBURY
; REGISTRATION NUMBER: 30,971
; REFERENCE/DOCKET NUMBER: BB-1071-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-773-0164
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1488 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-108-010B-11

Query Match 4.8%; Score 17; DB 4; Length 1488;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 261 CCCCGAAGCAAG 277
Db 898 CCCCGAAGCAAG 914

Search completed: November 6, 2002, 18:13:18
Job time : 78 secs

THIS PAGE BLANK (USPTO)

Sequence 2, Application US/08795473B
Patent No. 6217858
GENERAL INFORMATION:
APPLICANT: Galun, Elithan
APPLICANT: Nahot, Orit
TITLE OF INVENTION: A Pharmaceutical Composition for Treating
TITLE OF INVENTION: Hepatitis B Virus (HBV) Infection
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Davidson, Davidson and Kappel, LLC
STREET: 1140 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS-DOS EDITOR
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/795,473B
FILING DATE: 11-FEB-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Davidson, Clifford M.
REGISTRATION NUMBER: 32,728
REFERENCE/DOCKET NUMBER: 963.1007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)-997-1028
TELEFAX: (212)-997-1037
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3319 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
US-08-795-473B-2

Query Match 5.1%; Score 18; DB 4; Length 3319;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 CCAGCATCAGTGTCTCAT 172
|||||
DB 1256 CCAGCATCAGTGTCTCAT 1273

RESULT 11
US-09-439-856-2
Sequence 2, Application US/09439856
Patent No. 6410009
GENERAL INFORMATION:
APPLICANT: Galun, Elithan
APPLICANT: Nahot, Orit
TITLE OF INVENTION: A Pharmaceutical Composition for Treating
TITLE OF INVENTION: Hepatitis B Virus (HBV) Infection
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Davidson, Davidson and Kappel, LLC
STREET: 1140 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS-DOS EDITOR
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/439,856
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/795,473
FILING DATE: 11-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Davidson, Clifford M.
REGISTRATION NUMBER: 32,728
REFERENCE/DOCKET NUMBER: 963.1007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)-997-1028
TELEFAX: (212)-997-1037
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3319 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
US-09-439-856-2

Query Match 5.1%; Score 18; DB 4; Length 3319;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 CCAGCATCAGTGTCTCAT 172
|||||
DB 1256 CCAGCATCAGTGTCTCAT 1273

RESULT 12
US-08-642-684-13
Sequence 13, Application US/08642684
Patent No. 5834253
GENERAL INFORMATION:
APPLICANT: HONG, GUO FAN
APPLICANT: FENG, ZHAI
TITLE OF INVENTION: A NEW DNA POLYMERASE WITH PROOF-READING
TITLE OF INVENTION: 3'-5' EXONUCLEASE ACTIVITY
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/642,684
FILING DATE: 03-MAY-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CHAPIN, MARIANA K.
REGISTRATION NUMBER: 35,843
REFERENCE/DOCKET NUMBER: 4694/219502
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3711
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 47 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-642-684-13

NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Davidson, Davidson and Kappel, LLC
STREET: 1140 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS-DOS EDITOR
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/795,473B
FILING DATE: 11-FEB-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Davidson, Clifford M.
REGISTRATION NUMBER: 963.1007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)-997-1037
TELEFAX: (212)-997-1037
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1486 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
US-08-795-473B-3

Query Match
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 155 CCAGCATCAGTGTGTCAT 172
Db 870 CCAGCATCAGTGTGTCAT 887

RESULT 7
US-09-439-856-3
Sequence 3, Application US/09439856
Patent No. 641009
GENERAL INFORMATION:
APPLICANT: Galun, Etchan
APPLICANT: Nahot, Orit
TITLE OF INVENTION: A Pharmaceutical Composition for Treating
TITLE OF INVENTION: Hepatitis B Virus (HBV) Infection
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Davidson, Davidson and Kappel, LLC
STREET: 1140 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MS-DOS EDITOR
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/439,856
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/795,473
FILING DATE: 11-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Davidson, Clifford M.

REGISTRATION NUMBER: 32,728
REFERENCE/DOCKET NUMBER: 963.1007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)-997-1028
TELEFAX: (212)-997-1037
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1486 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
US-09-439-856-3

Query Match
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 155 CCAGCATCAGTGTGTCAT 172
Db 870 CCAGCATCAGTGTGTCAT 887

RESULT 8
5171840-1
Patent No. 5171840
APPLICANT: KISHIMOTO, TADAMITSU
TITLE OF INVENTION: RECEPTOR PROTEIN FOR HUMAN B CELL
STIMULATORY FACTOR-2
NUMBER OF SEQUENCES: 11
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/298,694
FILING DATE: 19-JAN-1989
SEQ ID NO: 1
LENGTH: 2061

Query Match
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 155 CCAGCATCAGTGTGTCAT 172
Db 1065 CCAGCATCAGTGTGTCAT 1082

RESULT 9
5480796-1
Patent No. 5480796
APPLICANT: KISHIMOTO, TADAMITSU
TITLE OF INVENTION: ANTIBODIES AGAINST THE RECEPTOR PROTEIN
FOR HUMAN B CELL STIMULATORY FACTOR-2
NUMBER OF SEQUENCES: 8
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/907,650
FILING DATE: 02-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 298,694
FILING DATE: 19-JAN-1989
SEQ ID NO: 1
LENGTH: 2061

Query Match
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 155 CCAGCATCAGTGTGTCAT 172
Db 1065 CCAGCATCAGTGTGTCAT 1082

RESULT 10
US-08-795-473B-2

RESULT 2

US-09-334-220-4
; Sequence 4, Application US/09334220
; Patent No. 6323177
; GENERAL INFORMATION:
; APPLICANT: St. Jude's Children's Research Hospital
; APPLICANT: Curran, Thomas
; APPLICANT: D'Arcangelo, Gabriella
; TITLE OF INVENTION: INTERACTION OF REELIN WITH VERY LOW
; TITLE OF INVENTION: DENSITY LIPOPROTEIN (VLDL) RECEPTOR FOR SCREENING AND
; FILE REFERENCE: 2427/0704
; CURRENT APPLICATION NUMBER: US/09/334,220
; CURRENT FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO: 4
; LENGTH: 11580
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-334-220-4

Query Match 7.4%; Score 26; DB 4; Length 11580;
Best Local Similarity 100.0%; Pred. No. 0.00024;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 208 GAGAAATTGAGCCCTTCATGT 233
|||||

Db 1067 GAGAAATTGAGCCCTTCATGT 1092
|||||

RESULT 3

US-08-627-151A-15
; Sequence 15, Application US/08627151A
; Patent No. 5866341
; GENERAL INFORMATION:
; APPLICANT: SPINELLA, Dominic
; APPLICANT: BECHERER, Kathleen
; APPLICANT: BROWN, Steven
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: SCREENING DRUG LIBRARIES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gen-Probe Incorporated
; STREET: 10210 Genetic Center Drive
; City: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92121

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/627,151A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:

ATTORNEY/AGENT INFORMATION:
; NAME: Fisher, Carlos A
; REGISTRATION NUMBER: 36,510
; TELEPHONE: 619-410-8926
; TELEFAX: 619-410-8928
; TELEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1074 base pairs

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-627-151A-15

Query Match 5.1%; Score 18; DB 2; Length 1074;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 CCAGCATCACTGTGTCA 172
|||||

Db 819 CCAGCATCACTGTGTCA 836
|||||

RESULT 4

5171840-8
; Patent No. 5171840
; APPLICANT: KISHIMOTO, TADAMITSU
; TITLE OF INVENTION: RECEPTOR PROTEIN FOR HUMAN B CELL
; STIMULATORY FACTOR-2
; NUMBER OF SEQUENCES: 11
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/298,694
; FILING DATE: 19-JAN-1989
; SEQ ID NO: 8
; LENGTH: 1404
5171840-8

Query Match 5.1%; Score 18; DB 6; Length 1404;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 CCAGCATCACTGTGTCA 172
|||||

Db 819 CCAGCATCACTGTGTCA 836
|||||

RESULT 5

5480796-8
; Patent No. 5480796
; APPLICANT: KISHIMOTO, TADAMITSU
; TITLE OF INVENTION: ANTIBODIES AGAINST THE RECEPTOR PROTEIN
; FOR HUMAN B CELL STIMULATORY FACTOR-2
; NUMBER OF SEQUENCES: 8
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/907,650
; FILING DATE: 02-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 298,694
; FILING DATE: 19-JAN-1989
; SEQ ID NO: 8
; LENGTH: 1404
5480796-8

Query Match 5.1%; Score 18; DB 6; Length 1404;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 CCAGCATCACTGTGTCA 172
|||||

Db 819 CCAGCATCACTGTGTCA 836
|||||

RESULT 6

US-08-795-473B-3
; Sequence 3, Application US/08795473B
; Patent No. 6217858
; GENERAL INFORMATION:
; APPLICANT: Galun, Elhan
; APPLICANT: Nahot, Orit
; APPLICANT: Blum, Herbert E.
; TITLE OF INVENTION: A Pharmaceutical Composition for Treating
; Hepatitis B Virus (HBV) Infection

Query Match	100.0%	Score 351	DB 4	Length 11673
Best Local Similarity	100.0%	Pref. No. 1.4e-174		
Matches 351	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY 1	GAGCAGTGTGGCACCATCATGCATGGCAATGCTGTACCTCTTCTGTGAGCCGTAGGCGCTT	60		
Db 970	GAGCAGTGTGGCACCATCATGCATGGCAATGCTGTACCTCTTCTGTGAGCCGTAGGCGCTT	1029		
QY 61	CGAGAGCTGACCCACCACATGCTGAAACACACACAGATCTGTCTCCAGTTTTCATT	120		
Db 1030	CGAGAGCTGACCCACCACATGCTGAAACACACACAGATCTGTCTCCAGTTTTCATT	1089		
QY 121	GGGTGAGGATCATGTGATTTAGTTAGTACTGTGACCCACATCATCTGTGTATACGACCAAG	180		
Db 1090	GGGTGAGGATCATGTGATTTAGTTAGTACTGTGACCCACATCATCTGTGTATACGACCAAG	1149		
QY 181	AACAAATACCGCTGATTGGATTGCTGAGTGGAGAAATTTAGAGCCCTTCCAAATGTGAGCACA	240		
Db 1150	AACAAATACCGCTGATTGGATTGCTGAGTGGAGAAATTTAGAGCCCTTCCAAATGTGAGCACA	1209		
QY 241	GTCATCCACATCTCTTACTCTCCCGCAGGAAAGCCAAAGGGGAGAGCTGCAGTTCCAGTGG	300		
Db 1210	GTCATCCACATCTCTTACTCTCCCGCAGGAAAGCCAAAGGGGAGAGCTGCAGTTCCAGTGG	1269		
QY 301	AAACAGAGACGCTGCGAGTGGTGAAGTGTATGAGGCTCTCTGGGGCCMG	351		
Db 1270	AAACAGAGACGCTGCGAGTGGTGAAGTGTATGAGGCTCTCTGGGGCCMG	1320		